A Design Algorithm using External Perturbation to Improve Iterative Feedback Tuning Convergence

Huusom, Jakob Kjøbsted; Hjalmarsson, Håkan; Poulsen, Niels Kjølstad; Jørgensen, Sten Bay

Publication date:
2008

Document Version
Publisher's PDF, also known as Version of record

A Design Algorithm using External Perturbation to Improve Iterative Feedback Tuning Convergence

Jakob Kjøbsted Huusom[†], Håkan Hjalmarsson[⋆], Niels Kjølstad Poulsen[‡] and Sten Bay Jørgensen[†]

Technical Rapport PEC08-16 - August, 2008

† CAPEC @ Department of Chemical and Biochemical Engineering
Technical University of Denmark - DTU
Building 227. DK - 2800 Lyngby, Denmark
jkh@kt.dtu.dk, sjb@kt.dtu.dk

‡ Department of Informatics and Mathematical Modelling
Technical University of Denmark - DTU
Building 321. DK - 2800 Lyngby, Denmark
nkp@imm.dtu.dk

⋆ Automatic Control @ Department of Signals, Sensors and Systems
Royal Institute of Technology - KTH
Osquldag väg 10. SE - 100 44 Stockholm, Sweden
hakan.hjalmarsson@ee.kth.se

Abstract

Iterative Feedback Tuning constitutes an attractive control loop tuning method for processes in the absence of sufficient process insight. It is a purely data driven approach to optimization of the loop performance. The standard formulation ensures an unbiased estimate of the loop performance cost function gradient, which is used in a search algorithm for minimizing the performance cost.

A slow rate of convergence of the tuning method is often experienced when tuning for disturbance rejection. This is due to a poor signal to noise ratio in the process data. A method is proposed for increasing the data information content by introducing an optimal perturbation signal in the tuning algorithm. The perturbation signal design is based on a detailed analysis of the asymptotic accuracy of the tuning method. A formal algorithm for optimization of the perturbation signal spectrum when tuning for disturbance rejection is presented. Special cases where an explicit optimal design are available is discussed. The theoretical analysis is supported by a simulation example.

1 Introduction

Control design and tuning for disturbance rejection is one of the classical disciplines in control theory and control engineering science. Design of compensators for disturbance rejections is well documented (Åström, 1970; Box and Jenkins, 1970; Åström and Hägglund, 1995). Given a particular control design, the tuning of the control parameters can be conducted based on tuning rules or by minimization of some loop performance criterion. The performance criterion is typically a quadratic cost function with penalty on the process outputs and the control signals. Given a model of the system, the set of optimal control parameters which minimize the performance cost can be evaluated. In absence of a sufficiently reliable model, the tuning can be performed based on data obtained from the loop, by a data driven optimization. Iterative Feedback Tuning is a method for optimizing control parameters using closed loop data and this algorithm will form the basis for the modifications presented here. The basic algorithm was first presented in Hjalmarsson *et al.* (1994) and has since then been analyzed, extended and tested in a number of papers. Gevers (2002) and Hjalmarsson (2002) provide extensive overviews of the development of the method and references to applications.

The performance criterion, $F_N(y_t, u_t)$, used in the controller tuning is a function of the output and the control action for the control loop. Hence it is a function of the true system, the controller and external signals acting on the loop. We will use the set-up in Figure 1.1 where G is a causal scalar linear time-invariant system, C is the controller, which also is assumed to be causal scalar linear time-invariant, and where r_t is the reference signal and v_t is the disturbance, respectively. Assuming, as we will, that the disturbance is stochastic implies that the performance cost is itself a random variable. However, as in, e.g., LQG-control, it is natural to minimize the expected cost

$$F(\cdot) \triangleq \mathrm{E}\left[F_N(\cdot)\right] \tag{1.1}$$

where here $\mathrm{E}[\cdot]$ is the mathematical expectation over the random disturbances acting on the closed loop system. This notation will be used throughout this paper. Notice that in the following, when expectation of $F(\cdot)$ is taken, the expectation refers not to the random disturbances acting on the system when assessing the closed loop peformance. It refers to the random variables that have affected the experimental data that has been used to design the controller for which the performance of $F(\cdot)$ is to be assessed. In order words the expectation will be taken over the controller C which will be seen as a random variable.

Our objective is to design a controller such that F is minimized when $r_t \equiv 0$, i.e. we are

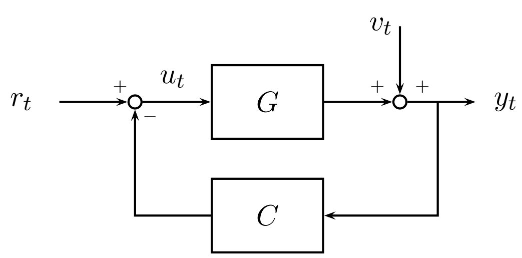


Figure 1.1: A general feedback loop designed for disturbance rejection. The process, G, and the compensator in the feedback loop, C, is given as scalar linear transfer functions.

interested in disturbance rejection. Adding a reference signal during the experimentation phase may however improve the quality of the obtained controller C. In Iterative Feedback Tuning, one tries to minimize F with respect to the controller using noisy closed loop experiments. The accuracy of this very much depends on the shape of the cost function F one tries to minimize. The sharper the optimum of F is, the easier it will be to find a good controller. Now, any change in the spectrum Φ_r of the reference signal, will affect the output spectrum Φ_y and the input spectrum Φ_u. Hence the reference signal spectrum affects the minimum and the shape of the performance cost surface. By designing the spectrum of an external reference it is consequently possible to shape the performance cost function in order to improve the convergence properties of the search algorithm in the tuning method for the control parameters. However, one has to bear in mind that shaping the cost function will also influence the location of the minimum in the controller parameter space. The cost function evaluated with external perturbation will be different from that of the original design problem when tuning for disturbance rejection. This is illustrated in Figure 1.2 where two examples of a quadratic cost function are shown as function of two control parameters. Let the original design F_0 refer to the disturbance rejection case where the reference signal to the loop is zero. F_1 is then the evaluation of the same cost function for the case with external perturbation where $\Phi_r \neq 0$. Since the contour lines of F_1 are closer together than for F_0, the optimization with the perturbation is less sensitive to the stochastic element in the evaluation of the performance cost. The price to be paid is that the method converges towards a different minimum. Despite this unfortunate consequence, successful simulation studies are reported with respect to convergence using Perturbed Iterative Feedback Tuning when tuning for disturbance rejection (Huusom *et al.*, 2008).

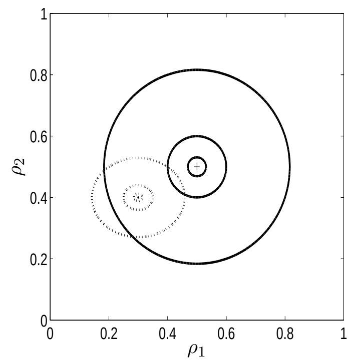


Figure 1.2: Contours and minima for two cost functions with equal levels for the contour lines. ρ_1 and ρ_2 are the control parameters. The full lines and the cross refer to the original design criterion, F_0. The dotted lines and the dot in the center is the cost function when affected by an external perturbation, F_1.

1.1 Formulating a design criterion

Let $F(\boldsymbol{\rho}, \boldsymbol{\vartheta})$ denote the cost function that we are interested in minimizing, where $\boldsymbol{\rho}$ and $\boldsymbol{\vartheta}$ represent the free control parameters which are to be tuned and a set of parameters which characterize the reference signal spectrum, respectively. The objective is to find the optimal $\boldsymbol{\rho}$ for a given $\boldsymbol{\vartheta} = \boldsymbol{\vartheta}^0$ which corresponds to $r_t \equiv 0$. We denote the optimum $\boldsymbol{\rho}$ by $\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta})$, indicating its dependence on $\boldsymbol{\vartheta}$. Since the system will be affected by noise it is only possible to obtain a minimizer, $\hat{\boldsymbol{\rho}}_n(\boldsymbol{\vartheta})$, with a certain accuracy; we use subscript n to denote that n iterations are performed in the tuning method. Hence Iterative Feedback Tuning will produce a solution with the following error

$$\boldsymbol{\Sigma}_n(\boldsymbol{\vartheta}) \triangleq \mathrm{E}\left[(\hat{\boldsymbol{\rho}}_n(\boldsymbol{\vartheta}) - \bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}))\,(\hat{\boldsymbol{\rho}}_n(\boldsymbol{\vartheta}) - \bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}))^T\right] \tag{1.2}$$

that has the property that it depends on $\boldsymbol{\vartheta}$. Using a continuity argument it may therefore be advantageous to optimize $\boldsymbol{\rho}$ for a $\boldsymbol{\vartheta} \neq \boldsymbol{\vartheta}^0$, i.e. it may be that the controller corresponding to $\boldsymbol{\vartheta}$ may result in a smaller expected cost for the desired excitation conditions (which correspond to $\boldsymbol{\vartheta}^0$) than the controller tuned with the desired operating conditions $\boldsymbol{\vartheta}^0$. This can be expressed as that it may hold that

$$\mathrm{E}\left[F(\hat{\boldsymbol{\rho}}_n(\boldsymbol{\vartheta}), \boldsymbol{\vartheta}^0)\right] < \mathrm{E}\left[F(\hat{\boldsymbol{\rho}}_n(\boldsymbol{\vartheta}^0), \boldsymbol{\vartheta}^0)\right] \tag{1.3}$$

Our objective is to determine operating conditions $\boldsymbol{\vartheta}$ such that $\mathrm{E}\left[F(\hat{\boldsymbol{\rho}}_n(\boldsymbol{\vartheta}), \boldsymbol{\vartheta}^0)\right]$ is minimized. This is a very difficult problem since $F(\hat{\boldsymbol{\rho}}_n(\boldsymbol{\vartheta}), \boldsymbol{\vartheta}^0)$ is a very complicated and non-linear function of the random disturbances originating from the experiments on which $\hat{\boldsymbol{\rho}}_n(\boldsymbol{\vartheta})$ is based. This in turn means that the expectation with respect to these random variables is very difficult to compute. Our approach to cope with this is to perform a local analysis, assuming $\boldsymbol{\vartheta}$ to be close to $\boldsymbol{\vartheta}^0$. Using Taylor expansion near the optimum we have that

$$\begin{aligned}
F(\hat{\boldsymbol{\rho}}_n(\boldsymbol{\vartheta}),\boldsymbol{\vartheta}^0) \approx{}& F(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}^0), \boldsymbol{\vartheta}^0) + \frac{\partial F(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}^0), \boldsymbol{\vartheta}^0)}{\partial \boldsymbol{\rho}}\left(\hat{\boldsymbol{\rho}}_n(\boldsymbol{\vartheta}) - \bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}^0)\right) + \\
& \frac{1}{2}\left(\hat{\boldsymbol{\rho}}_n(\boldsymbol{\vartheta}) - \bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}^0)\right)^T \frac{\partial^2 F(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}^0), \boldsymbol{\vartheta}^0)}{\partial \boldsymbol{\rho}^2}\left(\hat{\boldsymbol{\rho}}_n(\boldsymbol{\vartheta}) - \bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}^0)\right) \\
={}& F(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}^0), \boldsymbol{\vartheta}^0) + \frac{1}{2}\mathrm{Tr}\left\{\frac{\partial^2 F(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}^0), \boldsymbol{\vartheta}^0)}{\partial \boldsymbol{\rho}^2}\left(\hat{\boldsymbol{\rho}}_n(\boldsymbol{\vartheta}) - \bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}^0)\right)\left(\hat{\boldsymbol{\rho}}_n(\boldsymbol{\vartheta}) - \bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}^0)\right)^T\right\}
\end{aligned} \tag{1.4}$$

which means that

$$\begin{aligned}
\mathrm{E}&\left[F(\hat{\boldsymbol{\rho}}_n(\boldsymbol{\vartheta}), \boldsymbol{\vartheta}^0)\right] - F(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}^0), \boldsymbol{\vartheta}^0) \\
&\approx \frac{1}{2}\mathrm{Tr}\left\{\frac{\partial^2 F(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}^0), \boldsymbol{\vartheta}^0)}{\partial \boldsymbol{\rho}^2}\mathrm{E}\left[\left(\hat{\boldsymbol{\rho}}_n(\boldsymbol{\vartheta}) - \bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}^0)\right)\left(\hat{\boldsymbol{\rho}}_n(\boldsymbol{\vartheta}) - \bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}^0)\right)^T\right]\right\} \\
&= \frac{1}{2}\mathrm{Tr}\left\{\frac{\partial^2 F(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}^0), \boldsymbol{\vartheta}^0)}{\partial \boldsymbol{\rho}^2}\left(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}) - \bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}^0)\right)\left(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}) - \bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}^0)\right)^T\right\} + \\
&\quad \frac{1}{2}\mathrm{Tr}\left\{\frac{\partial^2 F(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}^0), \boldsymbol{\vartheta}^0)}{\partial \boldsymbol{\rho}^2}\boldsymbol{\Sigma}_n(\boldsymbol{\vartheta})\right\} \triangleq \Delta F_n(\boldsymbol{\vartheta})
\end{aligned} \tag{1.5}$$

Now, if $\boldsymbol{\Sigma}_n(\boldsymbol{\vartheta})$ can be evaluated then $\Delta F_n(\boldsymbol{\vartheta})$ is a quantity that can be minimized with respect to $\boldsymbol{\vartheta}$ in order to find the (approximately) optimal (reference) perturbation signal spectrum to be used in the experiments when tuning the controller parameters $\boldsymbol{\rho}$ using Iterative Feedback Tuning.

The two terms in $\Delta F_n(\boldsymbol{\vartheta})$ can be interpreted as follows: The first term is the bias error due to that $\boldsymbol{\vartheta} \neq \boldsymbol{\vartheta}^0$ is used in the optimization whereas the second term is the variance error incurred on $F(\hat{\boldsymbol{\rho}}_n(\boldsymbol{\vartheta}), \boldsymbol{\vartheta}^0)$. The bias error will typically increase as $\boldsymbol{\vartheta}$ moves away from $\boldsymbol{\vartheta}^0$. As noted above, it may be possible to decrease the variance error if $\boldsymbol{\vartheta}$ is suitably chosen. The optimal perturbation choice $\boldsymbol{\vartheta} = \bar{\boldsymbol{\vartheta}}$ will balance these two terms. The aim of this study is to construct a systematic and formal algorithm for designing an optimal external perturbation signal for Iterative Feedback Tuning of the disturbance rejection problem. Based on (1.5), this algorithm will minimize a design criterion which explicitly addresses this trade off between bias and variance error in the distribution of the n'th iterate in the tuning algorithm, $\boldsymbol{\rho}_n$.

The paper is organized as follows: Section 2 presents the basic Iterative Feedback Tuning algorithm for disturbance rejection. We also review an expression for the error $\boldsymbol{\Sigma}_n(\boldsymbol{\vartheta})$ of the method derived in Hildebrand *et al.* (2005b) for the disturbance rejection problem. In Section 3 the effect of adding an external perturbation signal to the loop in the tuning method is analyzed. This extends the result in Hildebrand *et al.* (2005b) and provides us with an expression for $\boldsymbol{\Sigma}_n(\boldsymbol{\vartheta})$ required for the computation of the expression on the right in (1.5). Then in Section 4, a formal design criterion for the perturbation spectrum is derived and a full algorithm, tuning for disturbance rejection with Perturbed Iterative Feedback Tuning using process insight, is constructed. Finally a simulation example serves to illustrate the advantages of introducing an optimal external perturbation signal in the tuning algorithm for the disturbance rejection case. Derivation of covariance expressions for the derivative of the performance cost function is given in an appendix.

2 Iterative Feedback Tuning for disturbance rejection

The algorithm for performing Iterative Feedback Tuning for disturbance rejection is illustrated in the following. The feedback loop in Figure 1.1 depicts the signals and transfer functions which will be used in the algorithm for tuning the parameters $\boldsymbol{\rho}$ in C. The objetive is to tuning the controller such that the effect of the noise, v_t, is rejected in an optimal sense.

The objective is to minimize the cost function:

$$F_N(\boldsymbol{\rho}_i) = \frac{1}{2N}\sum_{t=1}^{N}(y_t(\boldsymbol{\rho}_i) - y_t^d)^2 + \lambda(u_t(\boldsymbol{\rho}_i))^2 \tag{2.1}$$

where N number of data points in the discrete time horizon and y^d is the desired output response. For the disturbance rejection problem $r_t \equiv 0$ and hence $y_t^d = 0$. The sensitivity of the cost function with respect to the control parameters is

$$\boldsymbol{J}(\boldsymbol{\rho}_i) = \frac{\partial F_N(\boldsymbol{\rho}_i)}{\partial \boldsymbol{\rho}} = \frac{1}{N}\sum_{t=1}^{N} y_t(\boldsymbol{\rho}_i)\frac{\partial y_t(\boldsymbol{\rho}_i)}{\partial \boldsymbol{\rho}} + \lambda u_t(\boldsymbol{\rho}_i)\frac{\partial u_t(\boldsymbol{\rho}_i)}{\partial \boldsymbol{\rho}} \tag{2.2}$$

where

$$\frac{\partial y_t}{\partial \boldsymbol{\rho}}(\boldsymbol{\rho}) = -\frac{\partial C(\boldsymbol{\rho})}{\partial \boldsymbol{\rho}} G S^2(\boldsymbol{\rho}) v_t \tag{2.3}$$

$$\frac{\partial u_t}{\partial \boldsymbol{\rho}}(\boldsymbol{\rho}) = -\frac{\partial C(\boldsymbol{\rho})}{\partial \boldsymbol{\rho}} S^2(\boldsymbol{\rho}) v_t \tag{2.4}$$

The minimization of the cost function is realized by iterating in the gradient scheme

$$\boldsymbol{\rho}_{i+1} = \boldsymbol{\rho}_i - \gamma_i \mathbf{R}^{-1}\boldsymbol{J}(\boldsymbol{\rho}_i) \tag{2.5}$$

where $\mathbf{R}$ is a positive definite matrix. It could be chosen as the Hessian of the cost function with respect to the control parameters $\boldsymbol{\rho}$, or the identity matrix to achieve a Newton or a steepest decent algorithm respectively. If a model for the system is unknown, the gradients of the in- and output and hence the cost function gradient can not be evaluated analytically. An estimate of the performance cost function gradient is

$$\widehat{\boldsymbol{J}(\boldsymbol{\rho}_i)} = \frac{1}{N}\sum_{t=1}^{N} y_t(\boldsymbol{\rho}_i)\frac{\widehat{\partial y_t(\boldsymbol{\rho}_i)}}{\partial \boldsymbol{\rho}} + \lambda u_t(\boldsymbol{\rho}_i)\frac{\widehat{\partial u_t(\boldsymbol{\rho}_i)}}{\partial \boldsymbol{\rho}} \tag{2.6}$$

where $\frac{\widehat{\partial y_t(\boldsymbol{\rho}_i)}}{\partial \boldsymbol{\rho}}$ and $\frac{\widehat{\partial u_t(\boldsymbol{\rho}_i)}}{\partial \boldsymbol{\rho}}$ are estimates of (2.3) and (2.4) respectively. In the traditional Iterative Feedback Tuning framework the minimization of the cost function, (2.1), is based on data from two successive experiments (Hjalmarsson *et al.*, 1998).

- Collect data $\{y_t^1(\boldsymbol{\rho}_i), u_t^1(\boldsymbol{\rho}_i)\}_{t=1,..,N}$ where $r_t^1 = 0$
- Collect data $\{y_t^2(\boldsymbol{\rho}_i), u_t^2(\boldsymbol{\rho}_i)\}_{t=1,..,N}$ where $r_t^2 = -y_t^1$

This data is used to estimate the gradients of the in- and outputs

$$\widehat{\frac{\partial y_t}{\partial \boldsymbol{\rho}}} \triangleq \frac{\partial C(\boldsymbol{\rho}_i)}{\partial \boldsymbol{\rho}} y_t^2 \tag{2.7}$$

$$= \frac{\partial y_t}{\partial \boldsymbol{\rho}}(\boldsymbol{\rho}_i) + \frac{\partial C(\boldsymbol{\rho}_i)}{\partial \boldsymbol{\rho}} S(\boldsymbol{\rho}_i) v_t^2 \tag{2.8}$$

$$\widehat{\frac{\partial u_t}{\partial \boldsymbol{\rho}}} \triangleq \frac{\partial C(\boldsymbol{\rho}_i)}{\partial \boldsymbol{\rho}} u_t^2 \tag{2.9}$$

$$= \frac{\partial u_t}{\partial \boldsymbol{\rho}}(\boldsymbol{\rho}_i) - \frac{\partial C(\boldsymbol{\rho}_i)}{\partial \boldsymbol{\rho}} S(\boldsymbol{\rho}_i) C(\boldsymbol{\rho}_i) v_t^2 \tag{2.10}$$

where (2.7) and (2.9), are the estimators for the gradients of the in- and outputs. When these two expressions are used to form the estimate for the performance cost function gradient (2.6), (2.8) and (2.10) imply that the estimate can be split into two terms: An analytic bias term, S_N, and a variance term, E_N. The latter term is due to the noise present in the second experiment.

$$\widehat{\boldsymbol{J}(\boldsymbol{\rho}_i)} = S_N(\boldsymbol{\rho}_i) + E_N(\boldsymbol{\rho}_i) \tag{2.11}$$

where

$$\begin{aligned} S_N(\boldsymbol{\rho}) =& \frac{1}{N}\sum_{t=1}^{N}\left[y_t^1(\boldsymbol{\rho})\frac{\partial y_t(\boldsymbol{\rho})}{\partial \boldsymbol{\rho}} + \lambda u_t^1(\boldsymbol{\rho})\frac{\partial u_t(\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\right] \\ =& \frac{1}{N}\sum_{t=1}^{N}\left[(S(\boldsymbol{\rho})v_t^1)\left(-\frac{\partial C(\boldsymbol{\rho})}{\partial \boldsymbol{\rho}} G S(\boldsymbol{\rho})^2 v_t^1\right) + \lambda(-S(\boldsymbol{\rho})C(\boldsymbol{\rho})v_t^1)\left(-\frac{\partial C(\boldsymbol{\rho})}{\partial \boldsymbol{\rho}} S(\boldsymbol{\rho})^2 v_t^1\right)\right] \end{aligned} \tag{2.12}$$

$$\begin{aligned} E_N(\boldsymbol{\rho}) =& \frac{1}{N}\sum_{t=1}^{N}\left[y_t^1(\boldsymbol{\rho})\left(\frac{\widehat{\partial y_t(\boldsymbol{\rho})}}{\partial \boldsymbol{\rho}} - \frac{\partial y_t(\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\right) + \lambda u_t^1(\boldsymbol{\rho})\left(\frac{\widehat{\partial u_t(\boldsymbol{\rho})}}{\partial \boldsymbol{\rho}} - \frac{\partial u_t(\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\right)\right] \\ =& \frac{1}{N}\sum_{t=1}^{N}\left[(S(\boldsymbol{\rho})v_t^1)\left(\frac{\partial C(\boldsymbol{\rho})}{\partial \boldsymbol{\rho}} S(\boldsymbol{\rho}) v_t^2\right) + \lambda(-S(\boldsymbol{\rho})C(\boldsymbol{\rho})v_t^1)\left(-\frac{\partial C(\boldsymbol{\rho})}{\partial \boldsymbol{\rho}} S(\boldsymbol{\rho}) C(\boldsymbol{\rho}) v_t^2\right)\right] \end{aligned} \tag{2.13}$$

The expectation of the variance part is zero, since the noise signals from the first and second experiment are independent. The estimate of the cost function gradient produced by the Iterative Feedback Tuning method is therefore an unbiased realization.

Given that the noise v is a zero mean, weakly stationary random signal, the key contribution in Iterative Feedback Tuning, is that it supplies an unbiased estimate of the cost function gradient, without requiring a plant model estimate, $\hat{G}$, (Hjalmarsson *et al.*, 1998). Let the estimate, (2.6), be an unbiased and monotonically increasing function of $\boldsymbol{\rho}$. Using the estimate (2.6) in the gradient iteration (2.5) instead of the analytical expression (2.2), as a stochastic approximation method, will still make the algorithm converge to the expectation of the local minimizer provided that the sequence of γ_i in (2.5) fulfills (Robbins and Monro, 1951; Hildebrand *et al.*, 2003)

$$\sum_{i=1}^{\infty} \gamma_i^2 < \infty, \quad \sum_{i=1}^{\infty} \gamma_i = \infty. \tag{2.14}$$

This condition is fulfilled e.g. by having $\gamma_i = {}^{a}/_{i}$ where a is some positive constant.

A Gauss-Newton approximation of the Hessian to the performance cost function with respect to the controller parameters was suggested in Hjalmarsson *et al.* (1994). This first order approximation can be estimated using the available signals from the tuning method

$$\hat{\mathbf{H}} = \frac{1}{N} \sum_{t=1}^{N} \left[\widehat{\frac{\partial y_t}{\partial \boldsymbol{\rho}}} \left(\widehat{\frac{\partial y_t}{\partial \boldsymbol{\rho}}} \right)^T + \lambda \widehat{\frac{\partial u_t}{\partial \boldsymbol{\rho}}} \left(\widehat{\frac{\partial u_t}{\partial \boldsymbol{\rho}}} \right)^T \right] \tag{2.15}$$

This estimate will not be unbiased due to squared terms of the noise in the two experiments, but it will be positive definite. A modification (Solari and Gevers, 2004) that involves additional experiments in each iteration of the iterative Feedback Tuning algorithm produces an unbiased Hessian estimate.

2.1 Asymptotic accuracy of the tuning method

The stochastic contribution in the gradient estimate will affect the asymptotic convergence rate of the tuning method. A quantitative analysis was performed by Hildebrand *et al.* (2005a). The result is as follows: With n being the iteration number and $\bar{\boldsymbol{\rho}}$ the optimal set of parameters, the sequence of random variables, $\sqrt{n}(\boldsymbol{\rho}_n - \bar{\boldsymbol{\rho}})$, converge *in distribution* to a normally distributed random variable with zero mean and covariance matrix $\boldsymbol{\Sigma}$ according to

$$\begin{aligned} &\sqrt{n}(\boldsymbol{\rho}_n - \bar{\boldsymbol{\rho}}) \xrightarrow{D} \mathcal{N}(0, \boldsymbol{\Sigma}) \\ &\boldsymbol{\Sigma} = a^2 \int_0^{\infty} e^{\boldsymbol{A}t} \boldsymbol{R}^{-1} \mathrm{Cov} \left[\widehat{\boldsymbol{J}(\bar{\boldsymbol{\rho}})} \right] \boldsymbol{R}^{-1} e^{\boldsymbol{A}^T t} dt \end{aligned} \tag{2.16}$$

The result in (2.16) is valid given the following set of conditions hold:

1. The sequence $\boldsymbol{\rho}_n$ converges to a local isolated minimum $\bar{\boldsymbol{\rho}}$ of F
2. $\boldsymbol{H}(\bar{\boldsymbol{\rho}})$ is the true Hessian for $F(\boldsymbol{\rho})$ at $\bar{\boldsymbol{\rho}}$.
3. The gain sequence $\{\gamma_n\}$ in (2.5) is given by $\gamma_n = {}^{a}/_{n}$, where a is a positive constant.
4. There exists an index $\bar{n}$ and a matrix $\boldsymbol{R}$ such that $\boldsymbol{R}_n = \boldsymbol{R}$ for all $n > \bar{n}$.
5. The matrix $\boldsymbol{A} = {}^{1}/_{2}\boldsymbol{I} - a\boldsymbol{R}^{-1}\boldsymbol{H}(\bar{\boldsymbol{\rho}})$ is stable, i.e. the real part of all the eigenvalues is negative.
6. The covariance matrix $\mathrm{Cov}\left[\widehat{\boldsymbol{J}(\bar{\boldsymbol{\rho}})}\right]$ is positive definite.

The result in (2.16) means that asymptotically the distribution for the deviation between the n'th iterate of the controller parameter and the true optimum is known, and that the method converges to the true local minimizer of the performance cost function. In Hildebrand *et al.* (2005b) it is shown that the covariance expression for the distribution simplifies if $\boldsymbol{H}(\bar{\boldsymbol{\rho}})$, i.e. the true Hessian, is used as the matrix $\boldsymbol{R}$ in (2.5). Hence for a Newton-Raphson optimization

$$\boldsymbol{\Sigma} = \frac{a^2}{2a-1}\boldsymbol{R}^{-1}\mathrm{Cov}\left[\widehat{\boldsymbol{J}(\bar{\boldsymbol{\rho}})}\right]\boldsymbol{R}^{-1} \tag{2.17}$$

As a measure of the quality of the controller for a given iteration, n, in the tuning algorithm Hildebrand *et al.* (2005b) suggest the difference between the expected value of the performance cost with $C(\boldsymbol{\rho}_n)$ in the loop minus the theoretical minimum value. This quantity, ΔF_n, will be referred to as the *control quality index.*

$$\Delta F_n \triangleq \mathrm{E}[F(\boldsymbol{\rho}_n)] - F(\bar{\boldsymbol{\rho}}) \tag{2.18}$$

This index is by definition a positive measure. Expanding it in a Taylor series around the optimum up to second order gives the approximation:

$$\Delta F_n \approx \frac{1}{2}\mathrm{E}\left[\Delta\bar{\boldsymbol{\rho}}_n^T\boldsymbol{H}(\bar{\boldsymbol{\rho}})\Delta\bar{\boldsymbol{\rho}}_n\right] \tag{2.19}$$

where $\Delta\bar{\boldsymbol{\rho}}_n = \boldsymbol{\rho}_n - \bar{\boldsymbol{\rho}}$. The following asymptotic expression when $\mathbf{H}(\bar{\boldsymbol{\rho}})\mathbf{R}^{-1} = \mathbf{I}$ is given in Hildebrand *et al.* (2005b):

$$\lim_{n\to\infty} n\mathrm{E}\left[\Delta\bar{\boldsymbol{\rho}}_n^T\boldsymbol{H}(\bar{\boldsymbol{\rho}})\Delta\bar{\boldsymbol{\rho}}_n\right] = \frac{a^2}{2a-1}\mathrm{Tr}\left\{\mathrm{Cov}\left[\widehat{\boldsymbol{J}(\bar{\boldsymbol{\rho}})}\right][\boldsymbol{R}^{-1}]\right\} \tag{2.20}$$

From this analysis, it is seen that the covariance of the gradient estimate for the performance cost function influences both the asymptotic covariance of the distribution of $\Delta\bar{\boldsymbol{\rho}}_n$ and the control performance quality measure given the parameters $\boldsymbol{\rho}_n$. It is therefore of interest to decompose this covariance expression. Due to the independence of the signals v_t^1 and v_t^2, the covariance of the gradient estimate in Equation (2.11) can be divided into the following contributions.

$$\begin{aligned}\mathrm{Cov}\left[\widehat{\boldsymbol{J}(\boldsymbol{\rho})}\right] &= \mathrm{Cov}[S_N(\boldsymbol{\rho})] + \mathrm{E}\left[E_N(\boldsymbol{\rho})S_N(\boldsymbol{\rho})^T\right] + \mathrm{E}\left[E_N(\boldsymbol{\rho})S_N(\boldsymbol{\rho})^T\right]^T + \mathrm{Cov}[E_N(\boldsymbol{\rho})]\\ &= \mathrm{Cov}[S_N(\boldsymbol{\rho})] + \mathrm{Cov}[E_N(\boldsymbol{\rho})]\end{aligned} \tag{2.21}$$

Assuming that the disturbance $\{v_t\}$ is a Gaussian process, the asymptotic frequency-domain expressions of the two remaining terms are (Hildebrand *et al.*, 2005a):

$$\begin{aligned}\lim_{N\to\infty} N\mathrm{Cov}[S_N(\boldsymbol{\rho})] =& \frac{2}{2\pi}\int_{-\pi}^{\pi}|S(e^{j\omega},\boldsymbol{\rho})|^4\Phi_v^2(\omega)\times\\ &\mathcal{R}e\left\{[G(e^{j\omega},\boldsymbol{\rho}) - \lambda\overline{C}(e^{j\omega},\boldsymbol{\rho})]S(e^{j\omega},\boldsymbol{\rho})\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right\}\times\\ &\mathcal{R}e\left\{[G(e^{j\omega},\boldsymbol{\rho}) - \lambda\overline{C}(e^{j\omega},\boldsymbol{\rho})]S(e^{j\omega},\boldsymbol{\rho})\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right\}^T d\omega\end{aligned} \tag{2.22}$$

$$\begin{aligned}\lim_{N\to\infty} N\mathrm{Cov}[E_N(\boldsymbol{\rho})] =& \frac{1}{2\pi}\int_{-\pi}^{\pi}|S(e^{j\omega},\boldsymbol{\rho})|^4\left[1+\lambda|C(e^{j\omega},\boldsymbol{\rho})|^2\right]^2\times\\ &\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\frac{\partial C^*(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\Phi_v^2(\omega)d\omega\end{aligned} \tag{2.23}$$

where $\bar{C}$ is the complex conjugate and C^* is the complex conjugate transpose of C. A derivation is presented in Appendix A.

3 Introducing external perturbations in the tuning

It is desired to improve the convergence rate and the asymptotic accuracy of the Iterative Feedback Tuning method. To achive this, the signal to noise ratio in data used in the tuning method must be increased. An external perturbation signal will be used as reference in the first of the two experiments used in the tuning algorithm. The experiments are then defined as follows:

- Collect data $\{y_t^1(\boldsymbol{\rho}_i), u_t^1(\boldsymbol{\rho}_i)\}_{t=1,..,N}$ where $r_t^1 = r_t^p$
- Collect data $\{y_t^2(\boldsymbol{\rho}_i), u_t^2(\boldsymbol{\rho}_i)\}_{t=1,..,N}$ where $r_t^2 = -y_t^1$

where the external input r_t^p is characterized by the spectrum Φ_{r^p}. A discussion on using external perturbations in the Iterative Feedback Tuning algorithm and an introduction to Perturbed Iterative Feedback Tuning are given in Huusom *et al.* (2008). The implication of introducing the external perturbation signal on the convergence properties of the method will be elaborated in the following.

The implication on the gradient estimate of the cost function from including this extra signal is

$$\begin{aligned} S_N(\boldsymbol{\rho}_i) =& \frac{1}{N}\sum_{t=1}^{N}\Bigg[\left(S(\boldsymbol{\rho}_i)(Gr_t^p + v_t^1)\right)\left(-\frac{\partial C(\boldsymbol{\rho}_i)}{\partial \boldsymbol{\rho}} GS(\boldsymbol{\rho}_i)^2(Gr_t^p + v_t^1)\right) + \\ & \lambda S(\boldsymbol{\rho}_i)(r_t^p - C(\boldsymbol{\rho}_i)v_t^1)\left(-\frac{\partial C(\boldsymbol{\rho}_i)}{\partial \boldsymbol{\rho}} S(\boldsymbol{\rho}_i)^2(Gr_t^p + v_t^1)\right)\Bigg] \qquad (3.1) \\ E_N(\boldsymbol{\rho}_i) =& \frac{1}{N}\sum_{t=1}^{N}\Bigg[\left(S(\boldsymbol{\rho}_i)(Gr_t^p + v_t^1)\right)\left(\frac{\partial C(\boldsymbol{\rho}_i)}{\partial \boldsymbol{\rho}} S(\boldsymbol{\rho}_i)v_t^2\right) + \\ & \lambda S(\boldsymbol{\rho}_i)(r_t^p - C(\boldsymbol{\rho}_i)v_t^1)\left(-\frac{\partial C(\boldsymbol{\rho}_i)}{\partial \boldsymbol{\rho}} C(\boldsymbol{\rho}_i)S(\boldsymbol{\rho}_i)v_t^2\right)\Bigg] \qquad (3.2) \end{aligned}$$

Given the following two complex functions

$$\Psi(e^{j\omega}, \boldsymbol{\rho}) = [G(e^{j\omega}, \boldsymbol{\rho}) - \lambda\overline{C}(e^{j\omega}, \boldsymbol{\rho})]S(e^{j\omega}, \boldsymbol{\rho})\frac{\partial C(e^{j\omega}, \boldsymbol{\rho})}{\partial \boldsymbol{\rho}} \qquad (3.3)$$

$$\Upsilon(e^{j\omega}, \boldsymbol{\rho}) = [|G(e^{j\omega}, \boldsymbol{\rho})|^2 + \lambda]S(e^{j\omega}, \boldsymbol{\rho})\frac{\partial C(e^{j\omega}, \boldsymbol{\rho})}{\partial \boldsymbol{\rho}} \qquad (3.4)$$

and assuming that the disturbance $\{v_t\}$ and the reference signal r_t are Gaussian processes, the asymptotic covariance expressions for $S_N(\boldsymbol{\rho})$ and $E_N(\boldsymbol{\rho})$ are given as (see Appendix B for details)

$$\begin{aligned} \lim_{N\to\infty} N\mathrm{Cov}[S_N(\boldsymbol{\rho})] = \frac{1}{2\pi}\int_{-\pi}^{\pi} |S(e^{j\omega}, \boldsymbol{\rho})|^4 \times \Bigg[\\ \mathcal{R}e\{G(e^{j\omega})\Upsilon(e^{j\omega}, \boldsymbol{\rho})\}\mathcal{R}e\{G(e^{j\omega})\Upsilon(e^{j\omega}, \boldsymbol{\rho})\}^T\Phi_{r^p}^2 + \mathcal{R}e\{\Psi(e^{j\omega}, \boldsymbol{\rho})\}\mathcal{R}e\{\Psi(e^{j\omega}, \boldsymbol{\rho})\}^T\Phi_v^2 + \\ \Bigg[2\mathcal{R}e\{G(e^{j\omega})\Upsilon(e^{j\omega}, \boldsymbol{\rho})\}\mathcal{R}e\{\Psi(e^{j\omega}, \boldsymbol{\rho})\}^T + \mathcal{I}m\{G(e^{j\omega})\Upsilon(e^{j\omega}, \boldsymbol{\rho})\}\mathcal{I}m\{\Psi(e^{j\omega}, \boldsymbol{\rho})\}^T - \\ \mathcal{I}m\{G(e^{j\omega})\Upsilon(e^{j\omega}, \boldsymbol{\rho})\}^T\mathcal{I}m\{\Psi(e^{j\omega}, \boldsymbol{\rho})\} + \mathcal{R}e\{G(e^{j\omega})\Psi(e^{j\omega}, \boldsymbol{\rho})\}\mathcal{R}e\{G(e^{j\omega})\Psi(e^{j\omega}, \boldsymbol{\rho})\}^T + \\ \mathcal{I}m\{G(e^{j\omega})\Psi(e^{j\omega}, \boldsymbol{\rho})\}\mathcal{I}m\{G(e^{j\omega})\Psi(e^{j\omega}, \boldsymbol{\rho})\}^T + \mathcal{R}e\{\Upsilon(e^{j\omega}, \boldsymbol{\rho})\}\mathcal{R}e\{\Upsilon(e^{j\omega}, \boldsymbol{\rho})\}^T + \\ \mathcal{I}m\{\Upsilon(e^{j\omega}, \boldsymbol{\rho})\}\mathcal{I}m\{\Upsilon(e^{j\omega}, \boldsymbol{\rho})\}^T\Bigg]\Phi_{r^p}\Phi_v\Bigg] d\omega \qquad (3.5) \end{aligned}$$

$$
\begin{aligned}
&\lim_{N\to\infty} N\mathrm{Cov}[E_N(\boldsymbol{\rho})] = \\
&\quad \frac{1}{2\pi}\int_{-\pi}^{\pi} |S(e^{j\omega},\boldsymbol{\rho})|^4 \left[1+\lambda|C(e^{j\omega},\boldsymbol{\rho})|^2\right]^2 \overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}} \left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T \Phi_v^2 + \\
&\quad \left[\mathcal{R}e\{\Psi(e^{j\omega},\boldsymbol{\rho}))\}\mathcal{R}e\{\Psi(e^{j\omega},\boldsymbol{\rho}))\}^T + \mathcal{I}m\{\Psi(e^{j\omega},\boldsymbol{\rho}))\}\mathcal{I}m\{\Psi(e^{j\omega},\boldsymbol{\rho}))\}^T\right] \times \\
&\quad |S(e^{j\omega},\boldsymbol{\rho})|^4 \Phi_{r^p}\Phi_v d\omega
\end{aligned} \tag{3.6}
$$

In relation to experimental design of the perturbation spectrum it is important to know how the Hessian approximation is affected:

$$
\hat{\mathbf{H}} = \frac{1}{N}\sum_{t=1}^{N}\left(\widehat{\frac{\partial y_t}{\partial\boldsymbol{\rho}}}\left(\widehat{\frac{\partial y_t}{\partial\boldsymbol{\rho}}}\right)^T + \lambda\widehat{\frac{\partial u_t}{\partial\boldsymbol{\rho}}}\left(\widehat{\frac{\partial u_t}{\partial\boldsymbol{\rho}}}\right)^T\right) \tag{3.7}
$$

where

$$
\widehat{\frac{\partial y_t}{\partial\boldsymbol{\rho}}} = \frac{\partial C(\boldsymbol{\rho}_i)}{\partial\boldsymbol{\rho}}\left(GS(\boldsymbol{\rho}_i)^2(Gr_t^p + v_t^1) + S(\boldsymbol{\rho}_i)v_t^2\right) \tag{3.8}
$$

$$
\widehat{\frac{\partial u_t}{\partial\boldsymbol{\rho}}} = \frac{\partial C(\boldsymbol{\rho}_i)}{\partial\boldsymbol{\rho}}\left(S(\boldsymbol{\rho}_i)^2(Gr_t^p + v_t^1) + C(\boldsymbol{\rho}_i)S(\boldsymbol{\rho}_i)v_t^2\right) \tag{3.9}
$$

hence

$$
\begin{aligned}
\hat{\mathbf{H}}(e^{j\omega}) =& \frac{1}{2\pi N}\int_{-\pi}^{\pi} \frac{\partial C(e^{j\omega},\boldsymbol{\rho}_i)}{\partial\boldsymbol{\rho}}\frac{\partial C^*(e^{j\omega},\boldsymbol{\rho}_i)}{\partial\boldsymbol{\rho}}\times \\
&\left(|G(e^{j\omega})|^2|S(e^{j\omega},\boldsymbol{\rho}_i)|^4\left(|G(e^{j\omega})|^2\Phi_{r^p}+\Phi_v\right) + |S(e^{j\omega},\boldsymbol{\rho}_i)|^2\Phi_v\right) + \\
&\lambda\frac{\partial C(e^{j\omega},\boldsymbol{\rho}_i)}{\partial\boldsymbol{\rho}}\frac{\partial C^*(e^{j\omega},\boldsymbol{\rho}_i)}{\partial\boldsymbol{\rho}}\times \\
&\left(|S(e^{j\omega},\boldsymbol{\rho}_i)|^4\left(|G(e^{j\omega})|^2\Phi_{r^p}+\Phi_v\right) + |C(e^{j\omega},\boldsymbol{\rho}_i)|^2|S(e^{j\omega},\boldsymbol{\rho}_i)|^2\Phi_v\right)d\omega
\end{aligned} \tag{3.10}
$$

$$
\begin{aligned}
=& \frac{1}{2\pi N}\int_{-\pi}^{\pi} \frac{\partial C(e^{j\omega},\boldsymbol{\rho}_i)}{\partial\boldsymbol{\rho}}\frac{\partial C^*(e^{j\omega},\boldsymbol{\rho}_i)}{\partial\boldsymbol{\rho}}\times \\
&\Big[(|G(e^{j\omega})|^2+\lambda)|S(e^{j\omega},\boldsymbol{\rho}_i)|^4(|G(e^{j\omega})|^2\Phi_{r^p}+\Phi_v) + \\
&(1+\lambda|C(e^{j\omega},\boldsymbol{\rho}_i)|^2)|S(e^{j\omega},\boldsymbol{\rho}_i)|^2\Phi_v\Big]d\omega
\end{aligned} \tag{3.11}
$$

From the expressions in this section, it can be seen how external perturbation will affect the relevant functions in relation to the covariance of the cost function gradient estimate.

- The asymptotic expressions for S_N and E_N are affine functions in the following variables. $S_N = f(\Phi_{r^p}^2, \Phi_v^2, \Phi_{r^p}\Phi_v)$ and $E_N = f(\Phi_v^2, \Phi_{r^p}\Phi_v)$, hence the asymptotic covariance estimate is also an affine function in $\Phi_{r^p}^2$, Φ_v^2 and $\Phi_{r^p}\Phi_v$.
- The Hessian estimate is an affine function in Φ_{r^p} and Φ_v only.

3.1 Unbiased gradient estimation with perturbation

From the general feedback loop, Figure 1.1, it is seen that the closed loop transfer functions are given by

$$y_t = GS(\boldsymbol{\rho}_i)r_t^p + S(\boldsymbol{\rho}_i)v_t \tag{3.12}$$

$$u_t = S(\boldsymbol{\rho}_i)r_t^p - C(\boldsymbol{\rho}_i)S(\boldsymbol{\rho}_i)v_t \tag{3.13}$$

It would be interesting to have a design of r^p which would not change the dynamics in the response of y or u with respect to the inputs, compared to the unperturbed case. If $r_t^p = \sqrt{\alpha}/G v_t$ would be realizable, the output in (3.12) will simplify to

$$y_t = GS(\boldsymbol{\rho}_i)\frac{\sqrt{\alpha}}{G}v_t + S(\boldsymbol{\rho}_i)v_t = (1+\sqrt{\alpha})S(\boldsymbol{\rho}_i)v_t$$

which is only a scaled expression of the output for the unperturbed case. This perturbation signal design will render the gradient estimate unbiased in case $\lambda = 0$ in (2.1), i.e. *minimum variance control*. It is optimal in the sense that this design will contribute to a better signal to noise ratio without driving the optimization of the control parameters to a biased optimum compared to the unperturbed case. In case where $r_p = \sqrt{\alpha}C(e^{j\omega}, \boldsymbol{\rho}_i)v_t$

$$u_t = S(\boldsymbol{\rho}_i)\sqrt{\alpha}C(e^{j\omega}, \boldsymbol{\rho}_i)v_t - C(\boldsymbol{\rho}_i)S(\boldsymbol{\rho}_i)v_t = (1+\sqrt{\alpha})C(\boldsymbol{\rho}_i)S(\boldsymbol{\rho}_i)v_t$$

which means that an equivalent design is possible with an unbiased gradient estimate, if the performance cost function only includes a penalty on the control (i.e. $\lambda \to \infty$). This is of course only of theoretical interest. The functional dependencies in (3.12) and (3.13) means that a perturbation design which will give scaled expressions for both y and u with respect to the unperturbed case does not exist.

In practical applications the actual random disturbance signal is unknown but the spectrum of the disturbance may be known. If the perturbation signal is generated using a signal with spectral properties equal to these of v, i.e. Φ_v, then the expected value of the gradient estimates will still be unbiased. If r_t^p and v_t are independent the spectrum of the output and the input in the two cases are:

$$\Phi_y = |G(e^{j\omega})|^2|S(e^{j\omega}, \boldsymbol{\rho}_i)|^2\Phi_{r^p} + |S(e^{j\omega}, \boldsymbol{\rho}_i)|^2\Phi_v \tag{3.14}$$

$$\Phi_u = |S(e^{j\omega}, \boldsymbol{\rho}_i)|^2\Phi_{r^p} + |C(e^{j\omega}, \boldsymbol{\rho}_i)|^2|S(e^{j\omega}, \boldsymbol{\rho}_i)|^2\Phi_v \tag{3.15}$$

Following the two optimal designs which has just be argued

$$\Phi_{r^p} = \frac{\alpha}{|G(e^{j\omega})|^2}\Phi_v \quad \Rightarrow \Phi_y = (1+\alpha)|S(e^{j\omega}, \boldsymbol{\rho}_i)|^2\Phi_v$$

$$\Phi_{r^p} = \alpha|C(e^{j\omega}, \boldsymbol{\rho}_i)|^2\Phi_v \Rightarrow \Phi_u = (1+\alpha)C(e^{j\omega}, \boldsymbol{\rho}_i)|S(e^{j\omega}, \boldsymbol{\rho}_i)|^2\Phi_v$$

From these expressions it is seen that the only requirement is knowledge of the noise spectrum and the magnitude functions $|G(e^{j\omega})|^2$ and $|C(e^{j\omega}, \boldsymbol{\rho}_i)|^2$ in order to produce a spectrum of the in- and output which are scaled with $(1+\alpha)$, compared to the unperturbed case. Insuring that the spectrum are scaled, is a less strict requirement than having the signals y and u scaled. E.g. let the true system model contain a time delay such that $G(q) = q^{-k}\bar{G}(q)$. Since $|G(e^{j\omega})|^2 = |\bar{G}(e^{j\omega})|^2$, a perturbation signal generated by $r_t^p = \sqrt{\alpha}/\bar{G} v_t$ would only scale Φ_y up by $(1+\alpha)$ but

$$y_t = GS(\boldsymbol{\rho}_i)\frac{\sqrt{\alpha}}{\bar{G}}v_t + S(\boldsymbol{\rho}_i)v_t = (1+\sqrt{\alpha}q^{-k})S(\boldsymbol{\rho}_i)v_t$$

which will change the dynamic response and hence render the gradient estimate of the minimum variance cost function biased. This result gives some information for generation of the optimal perturbation signal for disturbance rejection tuning of the minimum variance controller. It is desirable to have an input signal with the same spectral properties as the random disturbance acting on the system. Furthermore this signal will have to be filtered through the inverse of the true plant dynamics.

In practice it is not possible to generate an optimal perturbation signal since the plant dynamics is unknown. On the other hand, the analysis in this section offers an optimal design strategy for the perturbation signal in case a plant and noise covariance estimates are available.

3.2 Influence of the perturbation power

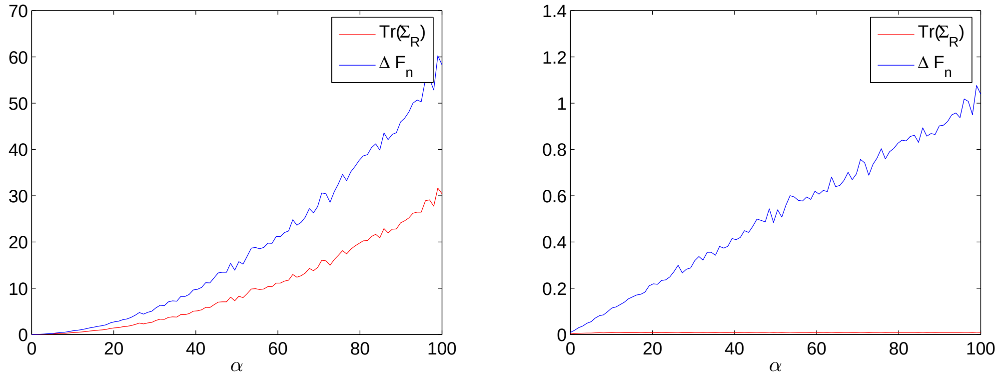
In this section it will be assumed that $\lambda = 0$ in the performance cost function which will therefore only depend on Φ_y. The perturbation signal spectrum will be chosen as $\Phi_{r^p} = (\alpha/|G(e^{j\omega})|^2)\Phi_v$ such that the only free parameter is α which will determine the power of the signal.

Using perturbations in the tuning algorithm will influence the covariance matrix of the performance cost function gradient estimate and hence the expected performance of the n'th iteration. Since expressions (3.5) and (3.6) show that the covariance matrix is proportional to the squared spectrum of the perturbation signal, it will be proportional to α^2. The true Hessian of the performance cost function, used in evaluation of Σ and ΔF_n, is independent of the perturbation, since this Hessian is evaluated at the optimum for the unperturbed problem. For $\alpha \to \infty$ it will therefore be expected that Σ and the control quality index will grow with α^2. In practice the true Hessian is not known and has to be estimated from the same perturbed data. Equation (3.11) shows that such a Hessian estimate is proportional to the perturbation spectrum and hence α. By substitution of the true Hessian with this perturbed Hessian estimate in the expressions for Σ and ΔF_n, it will be expected that Σ will approach a constant value when $\alpha \to \infty$ while the control quality index will grow linearly. These results are verified by simulation in Figure 3.1.

In case the perturbation signal is *kept constant* between iterations, the covariance expression for the performance cost will change. Since the perturbation signal does not change between iterations it will be regarded as a deterministic signal. Hence the multiplication between signals driven by the perturbation signal r_p will not contribute to the covariance. That implies that the term in S_N in (3.5) with the squared spectrum of the perturbation signal will be zero. Hence for a deterministic r_p

$$\begin{aligned}
\lim_{N\to\infty} N\mathrm{Cov}[S_N(\boldsymbol{\rho})] = \frac{1}{2\pi}\int_{-\pi}^{\pi} |S(e^{j\omega},\boldsymbol{\rho})|^4 \times \Bigg[& \left[\mathcal{R}e\{\Psi(e^{j\omega},\boldsymbol{\rho})\}\mathcal{R}e\{\Psi(e^{j\omega},\boldsymbol{\rho})\}^T\right]\Phi_v^2 + \\
& \Big[2\mathcal{R}e\{G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho})\}\mathcal{R}e\{\Psi(e^{j\omega},\boldsymbol{\rho})\}^T + \mathcal{I}m\{G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho})\}\mathcal{I}m\{\Psi(e^{j\omega},\boldsymbol{\rho})\}^T - \\
& \mathcal{I}m\{G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho})\}^T\mathcal{I}m\{\Psi(e^{j\omega},\boldsymbol{\rho})\} + \mathcal{R}e\{G(e^{j\omega})\Psi(e^{j\omega},\boldsymbol{\rho})\}\mathcal{R}e\{G(e^{j\omega})\Psi(e^{j\omega},\boldsymbol{\rho})\}^T + \\
& \mathcal{I}m\{G(e^{j\omega})\Psi(e^{j\omega},\boldsymbol{\rho})\}\mathcal{I}m\{G(e^{j\omega})\Psi(e^{j\omega},\boldsymbol{\rho})\}^T + \mathcal{R}e\{\Upsilon(e^{j\omega},\boldsymbol{\rho})\}\mathcal{R}e\{\Upsilon(e^{j\omega},\boldsymbol{\rho})\}^T + \\
& \mathcal{I}m\{\Upsilon(e^{j\omega},\boldsymbol{\rho})\}\mathcal{I}m\{\Upsilon(e^{j\omega},\boldsymbol{\rho})\}^T\Big]\Phi_{r^p}\Phi_v\Bigg]d\omega
\end{aligned} \tag{3.16}$$

The covariance expression for E_N in (3.6) remains unchanged. Having the same realization for the perturbation signal will give a covariance expression for the performance

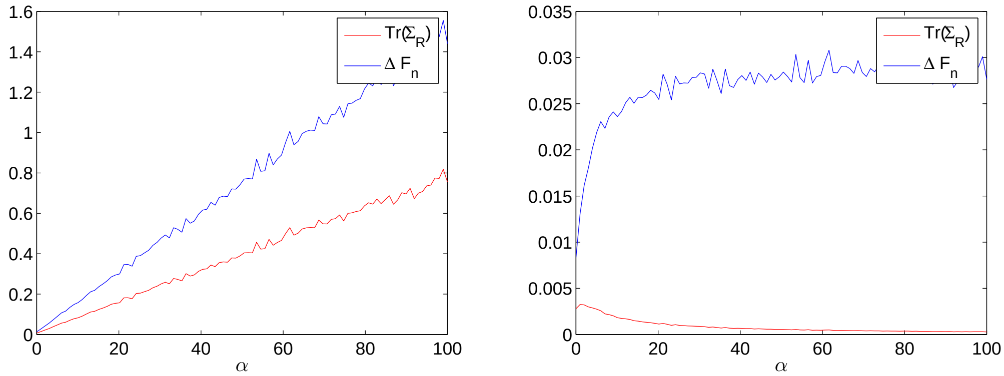


(a) True Hessian, $\mathbf{R} = \mathbf{H}$, for the unperturbed problem.

(b) Estimated Hessian, $\mathbf{R} = \hat{\mathbf{H}}$, using perturbations.

Figure 3.1: Simulations of $\mathrm{Tr}(\Sigma)$ and the control quality index, ΔF_n, for increasing power of the perturbation signal $\Phi_{r^p} = (\alpha/|G(e^{j\omega})|^2)\Phi_v$. 1000 data points are used in the simulation and 1000 repeated simulations are used to evaluate the covariance of the gradient estimate. The perturbation signal is *changed* between subsequent simulations.

cost gradient estimate which is proportional to the perturbation signal spectrum and not the spectrum squared. The influence of the deterministic perturbation signal on Σ and ΔF_n is shown in Figure 3.2. *It is seen that using a constant perturbation signal while the Hessian is estimated from data, produces a covariance matrix* Σ *which approaches zero as the power of the perturbation signal is increased.*



(a) True Hessian, $\mathbf{R} = \mathbf{H}$, for the unperturbed problem.

(b) Estimated Hessian, $\mathbf{R} = \hat{\mathbf{H}}$, using perturbations.

Figure 3.2: Simulations of $\mathrm{Tr}(\Sigma)$ and the control quality index, ΔF_n, for increasing power of the perturbation signal $\Phi_{r^p} = (\alpha/|G(e^{j\omega})|^2)\Phi_v$. 1000 data points are used in the simulation and 1000 repeated simulations are used to evaluate the covariance of the gradient estimate. The perturbation signal *remains unchanged* for all the subsequent simulations.

4 A formal design criterion for the perturbation spectrum

The previous section has shown that introducing an external perturbation signal in the first of the experiments in the Iterative Feedback Tuning algorithm, can improve the convergence and decrease the necessary number of iterations when the objective is disturbance rejection. In this section we summarize the formal design criterion, outlined in Section 1.1, and discuss practical issues.

Denoting the design variables of the reference spectrum by $\boldsymbol{\vartheta}$ (with $\boldsymbol{\vartheta}^0$ corresponding to a zero reference signal), we have from (1.5) that a suitable design criterion is

$$\begin{aligned}\Delta F_n(\boldsymbol{\vartheta}) \triangleq & \frac{1}{2}\mathrm{Tr}\left\{\frac{\partial^2 F(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}^0),\boldsymbol{\vartheta}^0)}{\partial \boldsymbol{\rho}^2}\left(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta})-\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}^0)\right)\left(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta})-\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}^0)\right)^T\right\}+ \\ & \frac{1}{2}\mathrm{Tr}\left\{\frac{\partial^2 F(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}^0),\boldsymbol{\vartheta}^0)}{\partial \boldsymbol{\rho}^2}\boldsymbol{\Sigma}_n(\boldsymbol{\vartheta})\right\}\end{aligned} \tag{4.1}$$

where

$$\boldsymbol{\Sigma}_n(\boldsymbol{\vartheta}) \triangleq \mathrm{E}\left[\left(\hat{\boldsymbol{\rho}}_n(\boldsymbol{\vartheta})-\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta})\right)\left(\hat{\boldsymbol{\rho}}_n(\boldsymbol{\vartheta})-\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta})\right)^T\right]$$

Recall that the first term in (4.1) is the bias, or the displacement of the optimal performance due to the external perturbation, and that the second term is the variance error. Under Conditions 1–6 in Section 2.1, (2.16) gives that $\boldsymbol{\Sigma}_n(\boldsymbol{\vartheta})$ can be expressed as

$$\boldsymbol{\Sigma}_n(\boldsymbol{\vartheta}) \approx \frac{1}{n}\boldsymbol{\Sigma}(\boldsymbol{\vartheta})$$

where n is the number of iterations that are going to be performed, and where

$$\boldsymbol{\Sigma}(\boldsymbol{\vartheta}) = a^2\int_0^\infty e^{\boldsymbol{A}t}\boldsymbol{R}^{-1}\mathrm{Cov}\left[\widehat{\boldsymbol{J}(\bar{\boldsymbol{\rho}},\boldsymbol{\vartheta})}\right]\boldsymbol{R}^{-1}e^{\boldsymbol{A}^Tt}dt \tag{4.2}$$

(recall that a is the gain in the step-size, i.e. at iteration n, the step-size is $\gamma_n = a/n$). In (4.2), $\mathrm{Cov}\left[\widehat{\boldsymbol{J}(\bar{\boldsymbol{\rho}},\boldsymbol{\vartheta})}\right]$ is given by (2.21)

$$\mathrm{Cov}\left[\widehat{\boldsymbol{J}(\boldsymbol{\rho},\boldsymbol{\vartheta})}\right] = \mathrm{Cov}[S_N(\boldsymbol{\rho},\boldsymbol{\vartheta})]+\mathrm{Cov}[E_N(\boldsymbol{\rho},\boldsymbol{\vartheta})] \tag{4.3}$$

where asymptotic (in the experiment length N) expressions for $\mathrm{Cov}[S_N(\boldsymbol{\rho},\boldsymbol{\vartheta})]$ and $\mathrm{Cov}[E_N(\boldsymbol{\rho}),\boldsymbol{\vartheta}]$ are given in (3.5)–(3.6) with Φ_{r^p} being the reference signal spectrum that corresponds to the parameter $\boldsymbol{\vartheta}$. Observe that these expressions hold when the disturbance v_t and the reference signal r_t are Gaussian distributed.

In case the gain direction $\mathbf{R}$ in the Iterative Feedback Tuning algorithm (2.5) is taken as $\frac{\partial^2 F(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}),\boldsymbol{\vartheta})}{\partial \boldsymbol{\rho}^2}$, the simplified expression (2.17) can be used resulting in that

$$\boldsymbol{\Sigma}_n(\boldsymbol{\vartheta}) \approx \frac{a^2}{n(2a-1)}\left[\frac{\partial^2 F(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}),\boldsymbol{\vartheta})}{\partial \boldsymbol{\rho}^2}\right]^{-1}\mathrm{Cov}\left[\widehat{\boldsymbol{J}(\bar{\boldsymbol{\rho}},\boldsymbol{\vartheta})}\right]\left[\frac{\partial^2 F(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}),\boldsymbol{\vartheta})}{\partial \boldsymbol{\rho}^2}\right]^{-1} \tag{4.4}$$

When full process knowledge is available all quantities in (4.1) can be computed from the equations above and thus one can optimize $\Delta F_n(\boldsymbol{\vartheta})$ in order to obtain a reference signal spectrum suitable for when using Iterative Feedback Tuning to tune a controller that is to be used for disturbance rejection. Since the design criterion $\Delta F_n(\boldsymbol{\vartheta})$ is based on a Taylor expansion it is recommended to introduce a constraint on the reference signal power in the optimization. There are many possibilities for parametrizing the reference spectrum. In the next section a straightforward method where filter coefficients are used as design variables $\boldsymbol{\vartheta}$. It is also possible to use a linear parametrization of the spectrum itself, we refer to Jansson and Hjalmarsson (2005) for details.

As in general experimental design algorithms, the evaluation of the optimal solution relies on knowledge of the true system which is not available (Goodwin and Payne, 1977; Gevers and Ljung, 1986; Bombois *et al.*, 2004). Therefore, practical use of the method will have to rely on an initial plant model. However, since the cost function appears to be smooth in many problems (see for example the next section), the accuracy of this model does not seem to be critical. The model may also be updated using the experimental data that is generated throughout the Iterative Feedback Tuning-experiments in order to successively improve the design.

5 An example

A simulation study is preformed in order to illustrate the ideas and advantages of introducing external perturbations in the Iterative Feedback Tuning method when tuning for disturbance rejection. For simplicity the control loop used is a discrete-time linear time-invariant transfer function model, and the controller has only two adjustable parameters. The random disturbance acting on the system is Gaussian white noise i.e. $e_t \in \mathcal{N}_{iid}(0, \sigma^2)$ where $\sigma = 1$. The nomenclature refers to the block diagram in Figure 1.1 where $v_t = H(q)e_t$.

$$\begin{aligned}
&\text{Plant model:} \quad && G(q) = \frac{q^{-1} - 0.5q^{-2}}{1 - 0.3q^{-1} - 0.28q^{-2}} \\
&\text{Noise model:} \quad && H(q) = \frac{1}{1 + 0.9q^{-1}} \\
&\text{Controller:} \quad && C(q) = \rho_1 + \rho_2 q^{-1}
\end{aligned} \tag{5.1}$$

This system was used in Hildebrand *et al.* (2005b) to test the advantages of optimal pre-filters in Iterative Feedback Tuning for disturbance rejection. The simulation study is divided into two cases. In the first case a minimum variance control design is used, hence $\lambda = 0$ in (2.1). In this case the optimal design of the perturbation signal is known analytically. The second case treats the more general case where penalty on both in- and outputs are included in the quadratic performance cost function. In the second case $\lambda = 0.25$ is chosen, and the optimal perturbation signal is designed by optimizing the parameters in a data filter.

Before we proceed, we remark that in this example we have replaced $\frac{\partial^2 F(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}^0), \boldsymbol{\vartheta}^0)}{\partial \boldsymbol{\rho}^2}$ in the second term of (4.1) by

$$\frac{F(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}), \boldsymbol{\vartheta}^0)}{F(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}), \boldsymbol{\vartheta})} \frac{\partial^2 F(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}), \boldsymbol{\vartheta})}{\partial \boldsymbol{\rho}^2} \tag{5.2}$$

This approximation is accurate when

$$\nu \triangleq \frac{d}{d\boldsymbol{\rho}} \left. \frac{\partial^2 F(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}), \boldsymbol{\vartheta})}{\partial \boldsymbol{\rho}^2} \right|_{\boldsymbol{\vartheta} = \boldsymbol{\vartheta}^0}$$

is small since $g(\boldsymbol{\vartheta})$ has first order derivative at $\boldsymbol{\vartheta} = \boldsymbol{\vartheta}^0$ given by $g'(\boldsymbol{\vartheta}^0) = \nu$.

$$g(\boldsymbol{\vartheta}) \triangleq \frac{F(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}), \boldsymbol{\vartheta}^0)}{F(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}), \boldsymbol{\vartheta})} \frac{\partial^2 F(\bar{\boldsymbol{\rho}}(\boldsymbol{\vartheta}), \boldsymbol{\vartheta})}{\partial \boldsymbol{\rho}^2}$$

5.1 Case one: Minimum variance control

In this section the external perturbation signal is given by $r_t^p = \sqrt{\alpha}/G v_t$ where the plant model G and the noise model H are assumed known. Hence α is the only free parameter which will give the input power of the perturbation signal. The tuning of the controller is performed for the minimum variance design where $\lambda = 0$ in the performance cost function (2.1). An external perturbation increases the value of the performance cost when applied. Figure 5.1 shows the normalized cost function as a surface on a grid of controller parameter values. These surface plots are smooth functions since the same noise realization has been used for each grid point and in both the perturbation signal design and in the evaluation of the cost function. The cost function is of course only a smooth function when the number of samples, $N \to \infty$, which is not practically realizable. In this simulation $N = 512$.

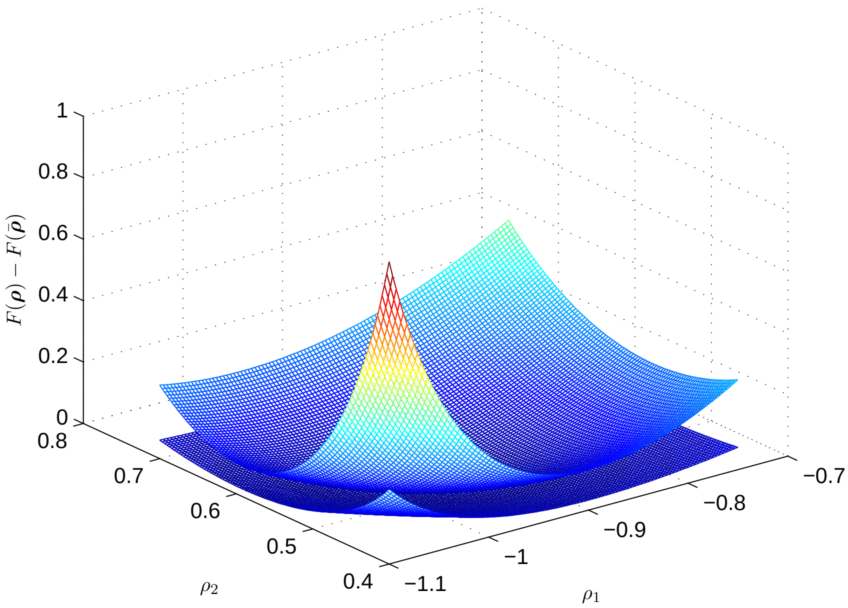
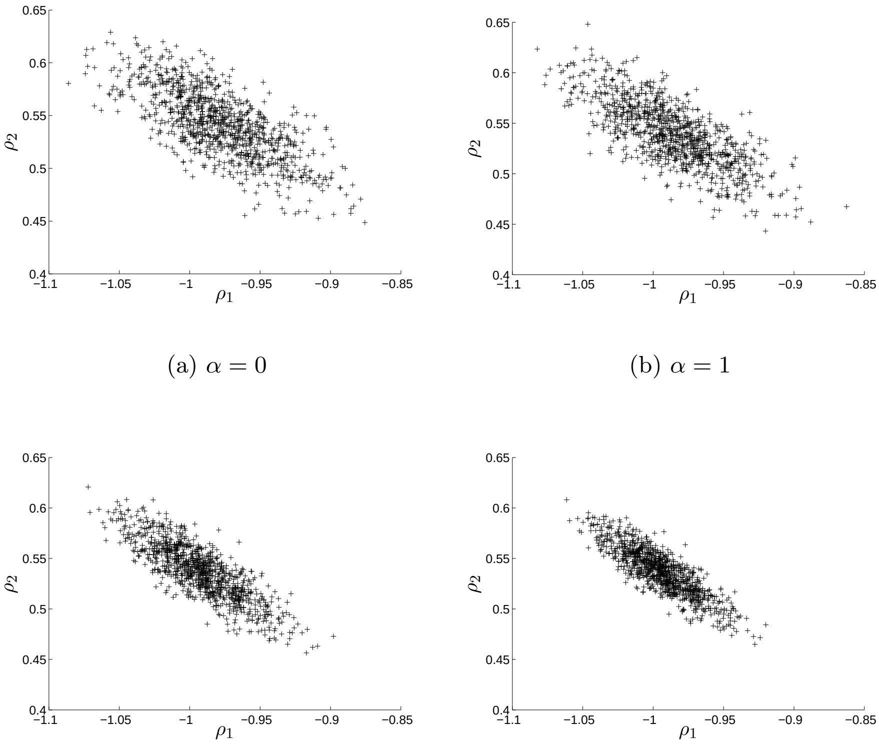


Figure 5.1: Surface plot of the normalized performance cost function on a control parameter grid. The lower surface is the performance cost when $\alpha = 0$ and the upper surface is for $\alpha = 1$. The same noise realization v_t has been used at each grid point and in the perturbation signal design in order to obtain a smooth surface.

The two surfaces have the same minimum in Figure 5.1, since for this idealized case

$$F(\boldsymbol{\rho}, \Phi_{r^p}) = (\sqrt{\alpha} + 1)^2 F(\boldsymbol{\rho}, 0)$$

This property means that the perturbation gives the desired change in the curvature of the performance cost function to yield a faster convergence. In order to illustrate this result further a series of Monte Carlo experiments are performed using Perturbed Iterative Feedback Tuning. Initially the control parameters has the optimal value, but due to the stochastic nature of the data the tuning will move the parameters away from this value for repeated iterations. In 1000 experiments, 10 iterations have been performed from the same optimal starting point, and the values of the resulting set of parameters has been saved. 1000 data points has been collected and used in each iteration of the tuning. For four different values of α in the perturbation signal, the results are presented in Figure 5.2. The variance and the cross-covariance of the resulting control parameters are reported in Table 5.1. From the results of the Monte Carlo



(a) $\alpha = 0$

(b) $\alpha = 1$

(c) $\alpha = 5$

(d) $\alpha = 10$

Figure 5.2: The final control parameters from 1000 Monte Carlo experiments each with 10 iterations in the Perturbed Iterative Feedback Tuning method. All iterations are initiated at the optimal value for the control parameters. The value of α in scaling the perturbation signal has been changed in four steps from zero up to 10.

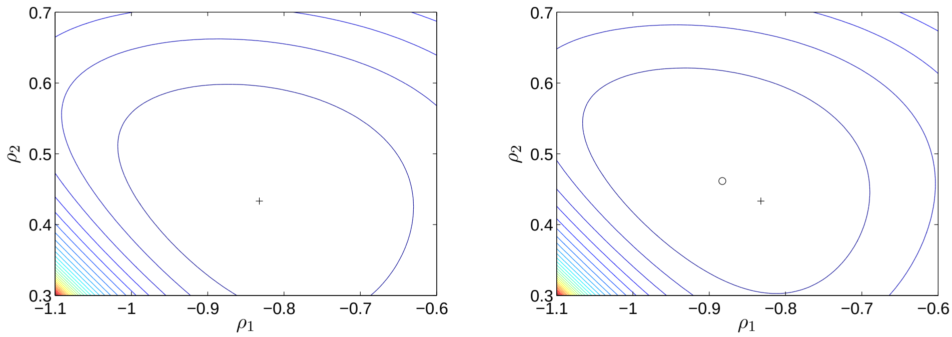
Variance	$\sigma^2_{\rho_1} \cdot 10^3$	$\sigma^2_{\rho_2} \cdot 10^3$	$\sigma_{\rho_1,\rho_2} \cdot 10^3$
$\alpha = 0$	1.24	1.03	-0.817
$\alpha = 1$	1.16	1.02	-0.820
$\alpha = 5$	0.757	0.743	-0.623
$\alpha = 10$	0.522	0.531	-0.451

Table 5.1: The variance and the cross-covariance for the resulting set of control parameters from 1000 Monte Carlo experiments each with 10 iterations in the Perturbed Iterative Feedback Tuning method. All iterations are initiated a t the optimal value for the control parameters. Results are given for different values of α.

simulations in Figure 5.2 and Table 5.1 it is obvious that increasing the value of α in the perturbation signal, produces an optimization problem with a statistically better defined optimum.

5.2 Case two: The general performance cost function

In this example the same system is used but the performance cost function is changed so that $\lambda = 0.25$. Initially the perturbation signal is formed in the same way as in the previous example, i.e. $r_t^p = \sqrt{\alpha}/G v_t$. Figure 5.3 show how the optimum of the cost function depends on the perturbation power when $\lambda \neq 0$. This figure also shows 30 contour lines for the cost functions, hence the contour lines in the two plots do not represent equal levels. The optimal control parameters for each surface and the corresponding value of the performance cost function are given in Table 5.2.



(a) $\alpha = 0$ (b) $\alpha = 1$

Figure 5.3: Contour plots with each 30 conture lines of the performance cost function, with $\lambda = 0.25$, on a control parameter grid. The perturbation signal $r_t^p = \sqrt{\alpha}/G v_t$ is applied and the optimal set of control parameters is marked with a + for $\alpha = 0$ and with $\circ$ for $\alpha = 1$. The same noise realization v_t has been used at each grid point and in the perturbation signal design in order to obtain a smooth surface.

	$\bar{\rho}_1$	$\bar{\rho}_2$	$F(\bar{\boldsymbol{\rho}})$	$F_{corr}(\bar{\boldsymbol{\rho}})$
$\alpha = 0$	-0.8323	0.4333	1.5008	-
$\alpha = 1$	-0.8828	0.4616	5.2697	1.5051

Table 5.2: The optimal set of control parameters for the two experiments with $\lambda = 0.25$ where the perturbation signal is given by $r_t^p = \sqrt{\alpha}/G v_t$. The value of the performance cost function for the optimal set is given together with the corrected value which compensates for the effect of perturbation.

In the general case where $\lambda \neq 0$ in the performance cost function, the ratio $\frac{F(\boldsymbol{\rho}(\Phi_{r^p}),0)}{F(\boldsymbol{\rho}(\Phi_{r^p}),\Phi_{r^p})}$ in (5.2) is not constant. This is evident from Figure 5.4 which also shows that the approximation is reasonable close to $\bar{\boldsymbol{\rho}}$. The figure also show that the curvature of this surface is small close to $\bar{\boldsymbol{\rho}}(\Phi_{r^p})$.

5.2.1 Optimizing the perturbation signal

Since the value for λ is not very large, it is possible that the optimal design for $\lambda = 0$ yields a reasonable filter choice. From Figure 5.3 it is seen that the optimum for the perturbed problem has not been moved very far from the unperturbed problem in the

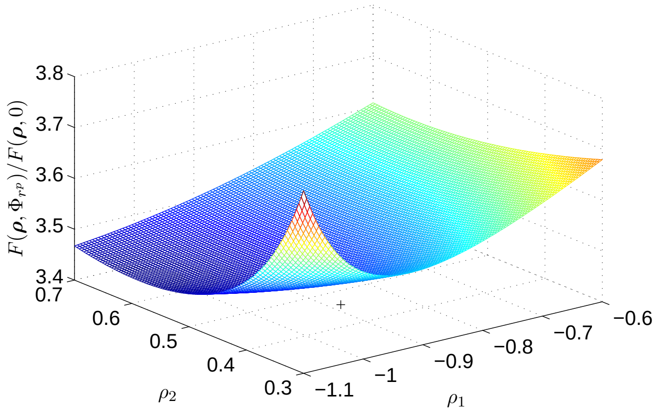


Figure 5.4: Surface plot of the performance cost function ratio $\frac{F(\boldsymbol{\rho}(\Phi_{r^p}),0)}{F(\boldsymbol{\rho}(\Phi_{r^p}),\Phi_{r^p})}$, with $\lambda = 0.25$, on a control parameter grid. The two evaluations of the performance cost are affected by perturbation signals $r_t^p = \sqrt{\alpha}/G v_t$ with $\alpha = 0$ and $\alpha = 1$ respectively. The same noise realization v_t has been used at each grid point and in the perturbation signal design in order to obtain a smooth surface. The point marked with + is the optimal set of control parameters for the perturbed problem.

control parameter space. Therefore the structure of the filter used to generate the perturbation signal is chosen identical to the inverted system model. The initial values for the filter parameters are selected as the model parameters $\boldsymbol{\theta}$.

$$r_t^p = \frac{1}{G(q)} H(q) e_t = G_{r^p}(q) H(q) e_t, \quad e_t \in \mathcal{N}_{iid}(0, \sigma^2) \tag{5.3}$$

where

$$G_{r^p}(q) = \frac{1 + \vartheta_1 q^{-1} + \vartheta_2 q^{-2}}{q^{-1} + \vartheta_3 q^{-2}}, \quad \boldsymbol{\vartheta}_0^T = \boldsymbol{\theta}^T = [-0.3 \ -0.28 \ -0.5] \tag{5.4}$$

hence non causal filtering is required. In this filter design the parameter α which adjusts the gain will not be included in $\boldsymbol{\vartheta}$ as a free parameter. Hence variance of the perturbation signal will be determined by the remaining free parameters. The optimal set of filter parameters can be detirmined by the minimization of the control quality index, $\Delta F_n(\Phi_{r^p})$ in (4.1) as an unconstrained problem. The reason being that the optimal perturbation power will be a trade off between the displacement of the optimal control parameters due to perturbation, and the distance between the expected and optimal performance. The optimal solution based on full process insight where computed as

$$\boldsymbol{\vartheta}_{opt}^T = [-8.115 \ -10.21 \ 0.5434]$$

and the control quality index were improved from $5.123 \cdot 10^{-3}$ to $0.2658 \cdot 10^{-3}$.

5.2.2 Perturbed Iterative Feedback Tuning

In the following four series of 1000 Monte Carlo experiments are performed each containing $n = 10$ iterations with Perturbed Iterative Feedback Tuning. Initially the loop starts with the optimal set of control parameters for the unperturbed operation.

- The first series is classical Iterative Feedback Tuning without a perturbation signal.

- In the second series the optimal designed perturbation filter for $\lambda = 0$ is used, hence $r_t^p = {}^{H(q)}/_{G(q)} e_t$. When λ is equal 0.25, this design is expected to produce a cloud of Monte Carlo solutions which is more dense but biased compeared to the first series.

- The third series uses the perturbation signal with the optimal parameters which was presented in Section 5.2.1. Since the variance of the perturbation signal is unconstrained in the optimization the variance for $r^p = G_{r^p}(q, \boldsymbol{\vartheta}_0)H(q)e_t$ is 2.842 while it is 165.6 for $r^p = G_{r^p}(q, \boldsymbol{\vartheta}_{opt})H(q)e_t$

- The fourth and last series, the third experiment is repeated but such a strong signal will not be allowed during the tuning and the filter is scaled accordingly. Using $\alpha G_{r^p}(q, \boldsymbol{\vartheta}_{opt})$ where $\alpha = \sqrt{2.842}/\sqrt{165.6}$, will give a variance of the optimal perturbation signal which is the same as for $r^p = G_{r^p}(q, \boldsymbol{\vartheta}_{opt})H(q)e_t$.

The results of these four trials are shown in Figure 5.5 as scatter plot of $\boldsymbol{\rho}_n$ together with the optimal solution for the unperturbed problem as a red square. Table 5.3 presents the mean value, the variance and the cross-covariance for the Monte Carlo solutions together with the optimal control parameter values for the unperturbed case.

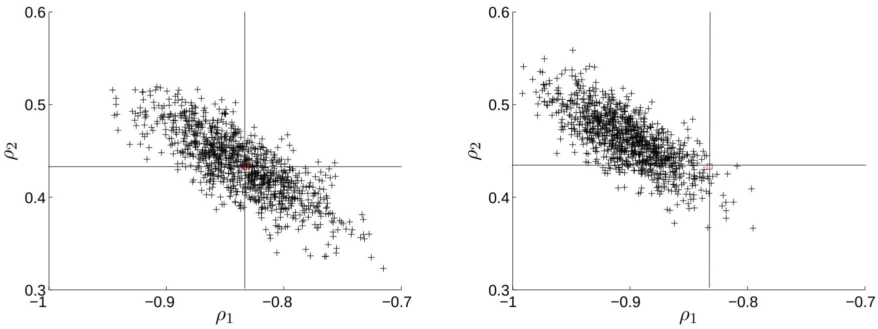
Statistic	$mean(\boldsymbol{\rho}_1)$	$mean(\boldsymbol{\rho}_2)$	$\sigma^2_{\rho_1} \cdot 10^3$	$\sigma^2_{\rho_2} \cdot 10^3$	$\sigma_{\rho_1,\rho_2} \cdot 10^3$
$r^p = 0$ (optimal)	-0.8323	0.4333	-	-	-
$r^p = 0$	-0.8369	0.4353	1.57	1.22	-1.09
$r^p = G_{r^p}(q, \boldsymbol{\vartheta}_0)H(q)e_t$	-0.9019	0.4628	1.07	0.899	-0.726
$r^p = G_{r^p}(q, \boldsymbol{\vartheta}_{opt})H(q)e_t$	-0.8382	0.4477	0.0275	0.0311	-0.0204
$r^p = \alpha G_{r^p}(q, \boldsymbol{\vartheta}_{opt})H(q)e_t$	-0.8371	0.4418	0.698	0.718	-0.488

Table 5.3: The mean, variance and the cross-covariance for the resulting set of control parameters from 1000 Monte Carlo experiments each with 10 iterations in the Perturbed Iterative Feedback Tuning method. All iterations are initiated at the optimal value for the control parameters for the unperturbed problem.

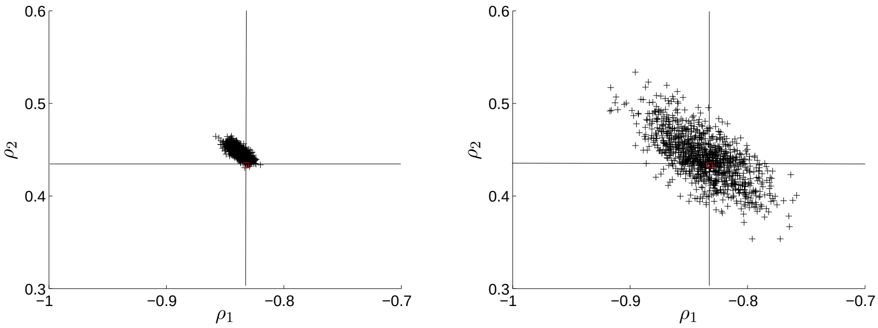
The results in Figure 5.5 and Table 5.3 clearly illustrates the advantage of introducing external perturbations when tuning for disturbance rejection. The optimal set of perturbation filter parameters both significantly reduce the variance of final control parameters from the 1000 Monte Carlo experiments, and reduce the displacement of the optimal control parameter solutions for the perturbed and unperturbed problem. It is possible to evaluate the optimal filter parameters for generating the perturbation signal as an unconstrained optimization. Constrains can then be included by a scaling the gain of the filter, which has been done for Figure 5.5d.

6 Conclusions

The convergence properties of the Perturbed Iterative Feedback Tuning algorithm for optimizing control parameters for disturbance rejection problems, have been investigated. Asymptotic expressions for the covariance of the cost function gradient have been derived and a control quality index for Perturbed Iterative Feedback Tuning is proposed. It is shown that using a deterministic external perturbation signal in the



(a) $r^p = 0$ (b) $r^p = G_{r^p}(q, \boldsymbol{\vartheta}_0)H(q)e_t$



(c) $r^p = G_{r^p}(q, \boldsymbol{\vartheta}_{opt})H(q)e_t$ (d) $r^p = \alpha G_{r^p}(q, \boldsymbol{\vartheta}_{opt})H(q)e_t$

Figure 5.5: The final control parameters from 1000 Monte Carlo experiments each with 10 iterations in the Perturbed Iterative Feedback Tuning method. All iterations are initiated at the optimal value for the control parameters for the unperturbed problem which is marked as a red square and with the straight lines. The perturbations signals used in series one to four corresponds to subfigure a, b, c, and d respectively.

tuning will affect the control quality index. The magnitude of the improvement depends on the power and the frequency content of the perturbation signal.

For minimal variance control design an analytical expression is derived for a parameterization of the perturbation signal which is optimal in the sense that it converges to the same set of control parameters as the unperturbed case. This optimal design is illustrated in a simulation example where it is shown that increasing the power of the perturbation signal improves the control quality index. For a general cost function with quadratic penalty on both the output and the input, there does not exist such an unbiased optimal parameterization of the perturbation signal. An algorithm for minimizing the control quality index for this general case is proposed based on process insight. This algorithm is shown to be able to produce a perturbation signal which significantly improves the control quality index. Hence Perturbed Iterative Feedback Tuning performs better than classical Iterative Feedback Tuning when tuning for disturbance rejection.

Investigation on optimal parameterization of the perturbation signal and the performance of Perturbed Iterative Feedback Tuning in case of very limited process insight remain for future work.

A Derivation of covariance expression for the cost function estimate $\hat{F}(\boldsymbol{\rho})$ in standard Iterative Feedback Tuning.

In this section the asymptotic covariance expression for the performance cost function estimate will be derived. This covariance expression has been shown in Section 2.1 to consist of the sum of the asymptotic covariance for the sums S_N and E_N reflecting the deterministic and the variance part of the gradient estimate respectively. These two sums are given by:

$$S_N(\boldsymbol{\rho}) = \frac{1}{N}\sum_{t=1}^{N}\left[(S(\boldsymbol{\rho})v_t^1)\left(-\frac{\partial C(\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}GS(\boldsymbol{\rho})^2 v_t^1\right) + \lambda(-S(\boldsymbol{\rho})C(\boldsymbol{\rho})v_t^1)\left(-\frac{\partial C(\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}S(\boldsymbol{\rho})^2 v_t^1\right)\right]$$

$$E_N(\boldsymbol{\rho}) = \frac{1}{N}\sum_{t=1}^{N}\left[(S(\boldsymbol{\rho})v_t^1)\left(\frac{\partial C(\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}S(\boldsymbol{\rho})v_t^2\right) + \lambda(-S(\boldsymbol{\rho})C(\boldsymbol{\rho})v_t^1)\left(-\frac{\partial C(\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}S(\boldsymbol{\rho})C(\boldsymbol{\rho})v_t^2\right)\right]$$

In the following derivations it will be used that S_N consists of four signals that are driven by the same noise realization. Hence all signals are correlated. E_N also consists of four signals but these are driven by two different realizations v_t^1 and v_t^2 from the same distribution.

Derivation A.1 (Covariance expressions for S_N) *Let Q_N (which is a generalization of the structure of S_N) be given by*

$$Q_N = \frac{1}{N}\sum_{t=1}^{N}[a(t)b(t) + c(t)d(t)] \tag{A.1}$$

where $a(t), b(t), c(t)$ and $d(t)$ are signals generated by filtering the Gaussian white noise signal $e(t)$ through the stable scalar transfer functions A and C and the vectors of stable transfer functions B and D. Hence

$$a(t) = Ae(t), \quad b(t) = Be(t), \quad c(t) = Ce(t), \quad d(t) = De(t)$$

Since all signals from $a(t)$ to $d(t)$ are correlated due to $e(t)$, one obtains:

$$Cov[Q_N] = E[Q_N Q_N^T] - E[Q_N]E[Q_N]^T \tag{A.2}$$

Evaluation of the first term gives

$$\begin{aligned}
E[Q_N Q_N^T] =& E\left[\frac{1}{N^2}\sum_{t=1}^{N}[a(t)b(t) + c(t)d(t)]\left(\sum_{t=1}^{N}a(t)b(t) + c(t)d(t)\right)^T\right] \\
=& \frac{1}{N^2}E\left[\sum_{t=1}^{N}[a(t)b(t) + c(t)d(t)]\left(\sum_{s=1}^{N}a(s)b(s) + c(s)d(s)\right)^T\right] \\
=& \frac{1}{N^2}E\left[\sum_{t,s=1}^{N}a(t)b(t)a(s)b(s)^T + a(t)b(t)c(s)d(s)^T + \right. \\
& \left. c(t)d(t)a(s)b(s)^T + c(t)d(t)c(s)d(s)^T\right] \\
=& \frac{1}{N^2}\left(E[\sum_{t,s=1}^{N}a(t)b(t)a(s)b(s)^T] + E[\sum_{t,s=1}^{N}a(t)b(t)c(s)d(s)^T] + \right. \\
& \left. E[\sum_{t,s=1}^{N}c(t)d(t)a(s)b(s)^T] + E[\sum_{t,s=1}^{N}c(t)d(t)c(s)d(s)^T]\right)
\end{aligned}$$

Using the following formula which is correct for the given properties of $a(t), b(t), c(t)$ and $d(t)$ and where α, β, δ and γ are fixed delays.

$$\begin{aligned}\overline{E}[a(t-\alpha)b(t-\beta)c(t-\gamma)d(t-\delta)] =& R_{ab}(\beta-\alpha)R_{cd^T}(\delta-\gamma) + R_{ac}(\gamma-\alpha)R_{bd^T}(\delta-\beta)+ \\ & R_{bc}(\delta-\alpha)R_{ad^T}(\gamma-\beta)\end{aligned}$$

the expression can be written as

$$\begin{aligned}E[Q_N Q_N^T] =& \frac{1}{N^2}\Bigg(\sum_{t,s=1}^{N}(R_{ab}(0)R_{ab^T}(0) + R_{aa}(t-s)R_{bb^T}(t-s) + R_{ba}(t-s)R_{ab^T}(t-s))+ \\ & \sum_{t,s=1}^{N}(R_{ab}(0)R_{cd^T}(0) + R_{ac}(t-s)R_{bd^T}(t-s) + R_{bc}(t-s)R_{ad^T}(t-s))+ \\ & \sum_{t,s=1}^{N}(R_{cd}(0)R_{ab^T}(0) + R_{ca}(t-s)R_{db^T}(t-s) + R_{da}(t-s)R_{cb^T}(t-s))+ \\ & \sum_{t,s=1}^{N}(R_{cd}(0)R_{cd^T}(0) + R_{cc}(t-s)R_{dd^T}(t-s) + R_{dc}(t-s)R_{cd^T}(t-s))\Bigg)\end{aligned}$$

The second term in (A.2) *will take the same form but will only have a contribution different from zero when the lag $t-s=0$. Hence*

$$E[Q_N]E[Q_N]^T = R_{ab}(0)R_{ab^T}(0) + R_{ab}(0)R_{cd^T}(0) + R_{cd}(0)R_{ab^T}(0) + R_{cd}(0)R_{cd^T}(0)$$

which means that the covariance of Q_N simplifies to

$$\begin{aligned}Cov[Q_N] =& \frac{1}{N^2}\sum_{t,s=1}^{N} R_{aa}(t-s)R_{bb^T}(t-s) + R_{ba}(t-s)R_{ab^T}(t-s)+ \\ & R_{ac}(t-s)R_{bd^T}(t-s) + R_{bc}(t-s)R_{ad^T}(t-s) + R_{ca}(t-s)R_{db^T}(t-s)+ \\ & R_{da}(t-s)R_{cb^T}(t-s) + R_{cc}(t-s)R_{dd^T}(t-s) + R_{dc}(t-s)R_{cd^T}(t-s)\end{aligned}$$

or

$$\begin{aligned}Cov[Q_N] =& \frac{1}{N}\sum_{\tau=-(N-1)}^{N-1}\frac{N-|\tau|}{N}(R_{aa}(\tau)R_{bb^T}(\tau) + R_{ba}(\tau)R_{ab^T}(\tau) + R_{ac}(\tau)R_{bd^T}(\tau)+ \\ & R_{bc}(\tau)R_{ad^T}(\tau) + R_{ca}(\tau)R_{db^T}(\tau) + R_{da}(\tau)R_{cb^T}(\tau) + R_{cc}(\tau)R_{dd^T}(\tau) + R_{dc}(\tau)R_{cd^T}(\tau))\end{aligned}$$

by letting $N \to \infty$, using Kronecker's lemma and applying the formula

$$\sum_{\tau=-\infty}^{\infty} R_{ab}(\tau)R_{cd^T}(\tau) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\Phi_{ab}(\omega)\overline{\Phi}_{cd^T}(\omega)d\omega \tag{A.3}$$

the following asymptotic expression appears

$$\begin{aligned}\lim_{N\to\infty} N\,Cov[Q_N] =& \frac{1}{2\pi}\int_{-\pi}^{\pi}\Phi_{aa}(\omega)\overline{\Phi}_{bb^T}(\omega) + \Phi_{ba}(\omega)\overline{\Phi}_{ab^T}(\omega)+ \\ & \Phi_{ac}(\omega)\overline{\Phi}_{bd^T}(\omega) + \Phi_{bc}(\omega)\overline{\Phi}_{ad^T}(\omega) + \Phi_{ca}(\omega)\overline{\Phi}_{db^T}(\omega)+ \\ & \Phi_{da}(\omega)\overline{\Phi}_{cb^T}(\omega) + \Phi_{cc}(\omega)\overline{\Phi}_{dd^T}(\omega) + \Phi_{dc}(\omega)\overline{\Phi}_{cd^T}(\omega)d\omega\end{aligned} \tag{A.4}$$

where $\overline{x}$ is the complex conjugate of x.

Letting A, B, C and D refer to the transfer functions in Equation (2.13), *the cross spectra can be evaluated using Equation* (A.5) *where x^* is the complex conjugated transpose of x.*

$$\Phi_{pq} = PQ^*\Phi_e \tag{A.5}$$

where $p(t) = Pe(t)$ and $q(t) = Qe(t)$ is a set of signals produced by filtering the same noise signal, $e(t)$ through the stable transfer functions P and Q.

$$
\begin{aligned}
\Phi_{aa} =& |S(e^{j\omega}, \boldsymbol{\rho})|^2 \Phi_v \\
\Phi_{bb^T} =& |G(e^{j\omega})|^2 |S(e^{j\omega}, \boldsymbol{\rho})|^4 \frac{\partial C(e^{j\omega}, \boldsymbol{\rho})}{\partial \boldsymbol{\rho}} \frac{\partial C^*(e^{j\omega}, \boldsymbol{\rho})}{\partial \boldsymbol{\rho}} \Phi_v \\
\Phi_{cc} =& \lambda |C(e^{j\omega}, \boldsymbol{\rho})|^2 |S(e^{j\omega}, \boldsymbol{\rho})|^2 \Phi_v \\
\Phi_{dd^T} =& \lambda |S(e^{j\omega}, \boldsymbol{\rho})|^4 \frac{\partial C(e^{j\omega}, \boldsymbol{\rho})}{\partial \boldsymbol{\rho}} \frac{\partial C^*(e^{j\omega}, \boldsymbol{\rho})}{\partial \boldsymbol{\rho}} \Phi_v \\
\Phi_{ba} =& -G(e^{j\omega}) S(e^{j\omega}, \boldsymbol{\rho}) |S(e^{j\omega}, \boldsymbol{\rho})|^2 \frac{\partial C(e^{j\omega}, \boldsymbol{\rho})}{\partial \boldsymbol{\rho}} \Phi_v \\
\Phi_{ab^T} =& -G^*(e^{j\omega}) S^*(e^{j\omega}, \boldsymbol{\rho}) |S(e^{j\omega}, \boldsymbol{\rho})|^2 \frac{\partial C^*(e^{j\omega}, \boldsymbol{\rho})}{\partial \boldsymbol{\rho}} \Phi_v \\
\Phi_{ac} =& -\sqrt{\lambda} C^*(e^{j\omega}, \boldsymbol{\rho}) |S(e^{j\omega}, \boldsymbol{\rho})|^2 \Phi_v \\
\Phi_{bd^T} =& \sqrt{\lambda} G(e^{j\omega}) |S(e^{j\omega}, \boldsymbol{\rho})|^4 \frac{\partial C(e^{j\omega}, \boldsymbol{\rho})}{\partial \boldsymbol{\rho}} \frac{\partial C^*(e^{j\omega}, \boldsymbol{\rho})}{\partial \boldsymbol{\rho}} \Phi_v \\
\Phi_{bc} =& \sqrt{\lambda} G(e^{j\omega}) S(e^{j\omega}, \boldsymbol{\rho}) C^*(e^{j\omega}, \boldsymbol{\rho}) |S(e^{j\omega}, \boldsymbol{\rho})|^2 \frac{\partial C(e^{j\omega}, \boldsymbol{\rho})}{\partial \boldsymbol{\rho}} \Phi_v \\
\Phi_{ad^T} =& -\sqrt{\lambda} S^*(e^{j\omega}, \boldsymbol{\rho}) |S(e^{j\omega}, \boldsymbol{\rho})|^2 \frac{\partial C^*(e^{j\omega}, \boldsymbol{\rho})}{\partial \boldsymbol{\rho}} \Phi_v \\
\Phi_{ca} =& -\sqrt{\lambda} C(e^{j\omega}, \boldsymbol{\rho}) |S(e^{j\omega}, \boldsymbol{\rho})|^2 \Phi_v \\
\Phi_{db^T} =& \sqrt{\lambda} G^*(e^{j\omega}) |S(e^{j\omega}, \boldsymbol{\rho})|^4 \frac{\partial C(e^{j\omega}, \boldsymbol{\rho})}{\partial \boldsymbol{\rho}} \frac{\partial C^*(e^{j\omega}, \boldsymbol{\rho})}{\partial \boldsymbol{\rho}} \Phi_v \\
\Phi_{da} =& -\sqrt{\lambda} S(e^{j\omega}, \boldsymbol{\rho}) |S(e^{j\omega}, \boldsymbol{\rho})|^2 \frac{\partial C(e^{j\omega}, \boldsymbol{\rho})}{\partial \boldsymbol{\rho}} \Phi_v \\
\Phi_{cb^T} =& \sqrt{\lambda} C(e^{j\omega}, \boldsymbol{\rho}) G^*(e^{j\omega}) S^*(e^{j\omega}, \boldsymbol{\rho}) |S(e^{j\omega}, \boldsymbol{\rho})|^2 \frac{\partial C^*(e^{j\omega}, \boldsymbol{\rho})}{\partial \boldsymbol{\rho}} \Phi_v \\
\Phi_{dc} =& \lambda S(e^{j\omega}, \boldsymbol{\rho}) C^*(e^{j\omega}, \boldsymbol{\rho}) |S(e^{j\omega}, \boldsymbol{\rho})|^2 \frac{\partial C(e^{j\omega}, \boldsymbol{\rho})}{\partial \boldsymbol{\rho}} \Phi_v \\
\Phi_{cd^T} =& \lambda C(e^{j\omega}, \boldsymbol{\rho}) S^*(e^{j\omega}, \boldsymbol{\rho}) |S(e^{j\omega}, \boldsymbol{\rho})|^2 \frac{\partial C^*(e^{j\omega}, \boldsymbol{\rho})}{\partial \boldsymbol{\rho}} \Phi_v
\end{aligned}
$$

Hence

$$\begin{aligned}
\Phi_{aa}\overline{\Phi}_{bb^T} =&|S(e^{j\omega},\boldsymbol{\rho})|^2|G(e^{j\omega})|^2\overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T|S(e^{j\omega},\boldsymbol{\rho})|^4\Phi_v^2\\
\Phi_{ba}\overline{\Phi}_{ab^T} =&S^2(e^{j\omega},\boldsymbol{\rho})G^2(e^{j\omega})\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T|S(e^{j\omega},\boldsymbol{\rho})|^4\Phi_v^2\\
\Phi_{cc}\overline{\Phi}_{dd^T} =&\lambda^2|S(e^{j\omega},\boldsymbol{\rho})|^2|C(e^{j\omega},\boldsymbol{\rho})|^2\overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T|S(e^{j\omega},\boldsymbol{\rho})|^4\Phi_v^2\\
\Phi_{dc}\overline{\Phi}_{cd^T} =&\lambda^2S^2(e^{j\omega},\boldsymbol{\rho})\overline{C}^2(e^{j\omega},\boldsymbol{\rho})\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T|S(e^{j\omega},\boldsymbol{\rho})|^4\Phi_v^2\\
\Phi_{ac}\overline{\Phi}_{bd^T} =&-\lambda\overline{G}(e^{j\omega})\overline{C}(e^{j\omega},\boldsymbol{\rho})|S(e^{j\omega},\boldsymbol{\rho})|^2\overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T|S(e^{j\omega},\boldsymbol{\rho})|^4\Phi_v^2\\
\Phi_{bc}\overline{\Phi}_{ad^T} =&-\lambda G(e^{j\omega})\overline{C}(e^{j\omega},\boldsymbol{\rho})S^2(e^{j\omega},\boldsymbol{\rho})\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T|S(e^{j\omega},\boldsymbol{\rho})|^4\Phi_v^2\\
\Phi_{ca}\overline{\Phi}_{db^T} =&-\lambda G(e^{j\omega})\overline{C}(e^{j\omega},\boldsymbol{\rho})|S(e^{j\omega},\boldsymbol{\rho})|^2\overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T|S(e^{j\omega},\boldsymbol{\rho})|^4\Phi_v^2\\
\Phi_{da}\overline{\Phi}_{cb^T} =&-\lambda G(e^{j\omega})\overline{C}(e^{j\omega},\boldsymbol{\rho})S^2(e^{j\omega},\boldsymbol{\rho})\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T|S(e^{j\omega},\boldsymbol{\rho})|^4\Phi_v^2
\end{aligned}$$

Inserting these expressions in (A.4) *one obtain after a few manipulations:*

$$\begin{aligned}
\lim_{N\to\infty} N\,Cov[S_N(\boldsymbol{\rho})] =&\frac{1}{2\pi}\int_{-\pi}^{\pi}|S(e^{j\omega},\boldsymbol{\rho})|^4\Phi_v^2\left[\Psi(e^{j\omega},\boldsymbol{\rho})\Psi(e^{j\omega},\boldsymbol{\rho})^T+\overline{\Psi}(e^{j\omega},\boldsymbol{\rho})\Psi(e^{j\omega},\boldsymbol{\rho})^T\right]d\omega\\
=&\frac{2}{2\pi}\int_{-\pi}^{\pi}|S(e^{j\omega},\boldsymbol{\rho})|^4\Phi_v^2\left[\mathcal{R}e\{\Psi(e^{j\omega},\boldsymbol{\rho})\}\mathcal{R}e\{\Psi(e^{j\omega},\boldsymbol{\rho})\}^T+\right.\\
&\left.j\mathcal{R}e\{\Psi(e^{j\omega},\boldsymbol{\rho})\}\mathcal{I}m\{\Psi(e^{j\omega},\boldsymbol{\rho})\}^T\right]d\omega
\end{aligned}$$

where

$$\Psi(e^{j\omega},\boldsymbol{\rho}) =[G(e^{j\omega},\boldsymbol{\rho})-\lambda\overline{C}(e^{j\omega},\boldsymbol{\rho})]S(e^{j\omega},\boldsymbol{\rho})\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}$$

Which is the same as in (2.23) *since the integration of* $2j\mathcal{R}e\{\Psi(e^{j\omega},\boldsymbol{\rho})\}\mathcal{I}m\{\Psi(e^{j\omega},\boldsymbol{\rho})\}^T$ *from* $-\pi$ *to* π *is zero for all* $\Psi(e^{j\omega},\boldsymbol{\rho})\in\mathbb{C}$. **q.e.d.**

Derivation A.2 (Covariance expressions for E_N) *For the derivation of the covariance of* E_N*, let the sum* Q_N *be a generalization of the structure of* E_N *according to*

$$Q_N=\frac{1}{N}\sum_{t=1}^{N}[a(t)b(t)+c(t)d(t)] \tag{A.6}$$

where $a(t), b(t), c(t)$ *and* $d(t)$ *are signals generated by filtering the two mutually independent white noise signals* $e(t)$ *and* $f(t)$ *through the stable scalar transfer functions* A *and* C *and the vectors of stable transfer functions* B *and* D*. Hence*

$$a(t)=Ae(t),\quad b(t)=Bf(t),\quad c(t)=Ce(t),\quad d(t)=Df(t)$$

Using the same derivation as above, but realizing that the cross correlation terms between signals driven by independent noise realizations will be zero, the covariance can be written as

$$\begin{aligned}
Cov[Q_N] =&\frac{1}{N^2}\sum_{t,s=1}^{N}R_{aa}(t-s)R_{bb^T}(t-s)+R_{ac}(t-s)R_{bd^T}(t-s)+\\
&R_{ca}(t-s)R_{db^T}(t-s)+R_{cc}(t-s)R_{dd^T}(t-s)
\end{aligned}$$

hence

$$\lim_{N\to\infty} N\,Cov[Q_N] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_{aa}(\omega)\overline{\Phi}_{bb^T}(\omega) + \Phi_{ac}(\omega)\overline{\Phi}_{bd^T}(\omega) + \Phi_{ca}(\omega)\overline{\Phi}_{db^T}(\omega) + \Phi_{cc}(\omega)\overline{\Phi}_{dd^T}(\omega)d\omega \tag{A.7}$$

When A, B, C and D refers to transfer functions in Equation (2.13) the cross spectra are

$$\begin{aligned}
\Phi_{aa} =& |S(e^{j\omega},\boldsymbol{\rho})|^2\Phi_v \\
\Phi_{cc} =& \lambda|C(e^{j\omega},\boldsymbol{\rho})|^2|S(e^{j\omega},\boldsymbol{\rho})|^2\Phi_v \\
\Phi_{ac} =& -\sqrt{\lambda}\overline{C}(e^{j\omega},\boldsymbol{\rho})|S(e^{j\omega},\boldsymbol{\rho})|^2\Phi_v \\
\Phi_{ca} =& -\sqrt{\lambda}C(e^{j\omega},\boldsymbol{\rho})|S(e^{j\omega},\boldsymbol{\rho})|^2\Phi_v \\
\Phi_{bb^T} =& |S(e^{j\omega},\boldsymbol{\rho})|^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\frac{\partial C^*(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\Phi_v \\
\Phi_{dd^T} =& \lambda|C(e^{j\omega},\boldsymbol{\rho})|^2|S(e^{j\omega},\boldsymbol{\rho})|^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\frac{\partial C^*(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\Phi_v \\
\Phi_{bd^T} =& -\sqrt{\lambda}\overline{C}(e^{j\omega},\boldsymbol{\rho})|S(e^{j\omega},\boldsymbol{\rho})|^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\frac{\partial C^*(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\Phi_v \\
\Phi_{db^T} =& -\sqrt{\lambda}C(e^{j\omega},\boldsymbol{\rho})|S(e^{j\omega},\boldsymbol{\rho})|^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\frac{\partial C^*(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\Phi_v
\end{aligned}$$

Hence

$$\begin{aligned}
\Phi_{aa}\overline{\Phi}_{bb^T} =& |S(e^{j\omega},\boldsymbol{\rho})|^4\overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T\Phi_v^2 \\
\Phi_{cc}\overline{\Phi}_{dd^T} =& \lambda^2|C(e^{j\omega},\boldsymbol{\rho})|^4|S(e^{j\omega},\boldsymbol{\rho})|^4\overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T\Phi_v^2 \\
\Phi_{ac}\overline{\Phi}_{bd^T} =& \lambda|C(e^{j\omega},\boldsymbol{\rho})|^2|S(e^{j\omega},\boldsymbol{\rho})|^4\overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T\Phi_v^2 \\
\Phi_{ca}\overline{\Phi}_{db^T} =& \Phi_{ac}\overline{\Phi}_{bd^T}
\end{aligned}$$

Inserting these expressions in (A.7) *gives:*

$$\lim_{N\to\infty} N\,Cov[E_N(\boldsymbol{\rho})] = \frac{1}{2\pi}\int_{-\pi}^{\pi}|S(e^{j\omega},\boldsymbol{\rho})|^4\left[1+\lambda|C(e^{j\omega},\boldsymbol{\rho})|^2\right]^2\times \overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T\Phi_v^2 d\omega$$

Since the result is a real number, this expression is the same as (2.23) *where the integrand is the complex conjugate of the expression above.* **q.e.d.**

B Derivation of covariance expression for the cost function estimate $\hat{F}(\boldsymbol{\rho})$ for Iterative Feedback Tuning with external perturbation.

In this section the asymptotic covariance expression for the performance cost function estimate will be derived for the case where the system is perturbed. As argued in Section 3 the covariance expression still consists only of the sum of the asymptotic covariances for the sums S_N and E_N which reflect the deterministic and the variance part of the gradient estimate respectively. These two sums are given by:

$$S_N(\boldsymbol{\rho}) = \frac{1}{N}\sum_{t=1}^{N}\Big[\left(S(\boldsymbol{\rho})(Gr_t^p + v_t^1)\right)\left(-\frac{\partial C(\boldsymbol{\rho})}{\partial \boldsymbol{\rho}} G S(\boldsymbol{\rho})^2 (Gr_t^p + v_t^1)\right) + \\ \lambda S(\boldsymbol{\rho})(r_t^p - C(\boldsymbol{\rho})v_t^1)\left(-\frac{\partial C(\boldsymbol{\rho})}{\partial \boldsymbol{\rho}} S(\boldsymbol{\rho})^2 (Gr_t^p + v_t^1)\right)\Big]$$

$$E_N(\boldsymbol{\rho}) = \frac{1}{N}\sum_{t=1}^{N}\Big[\left(S(\boldsymbol{\rho})(Gr_t^p + v_t^1)\right)\left(\frac{\partial C(\boldsymbol{\rho})}{\partial \boldsymbol{\rho}} S(\boldsymbol{\rho}) v_t^2\right) + \\ \lambda S(\boldsymbol{\rho})(r_t^p - C(\boldsymbol{\rho})v_t^1)\left(-\frac{\partial C(\boldsymbol{\rho})}{\partial \boldsymbol{\rho}} C(\boldsymbol{\rho}) S(\boldsymbol{\rho}) v_t^2\right)\Big]$$

In the following derivations r_t^p will be regarded as a signal driven by a white noise process with the same distribution as v_t^1 and v_t^2. r_t^p can be regarded as a deterministic signal which will be reused in every iteration of the Iterative Feedback tuning or alternatively a new signal could be generated for each iteration. The latter option will be assumed in the following. Hence S_N consists of eight signals arranged in two sets of four driven by r_t^p and v_t^1. E_N consists of six signals collected in three pairs each driven by r_t^p, v_t^1 or v_t^2.

Derivation B.1 (Covariance expressions for S_N) *For the derivation of the covariance of S_N, let the sum Q_N be a generalization of the structure of S_N according to*

$$Q_N = \frac{1}{N}\sum_{t=1}^{N}[(a1(t) + b1(t))(a2(t) + b2(t)) + (a3(t) + b3(t))(a4(t) + b4(t))] \\ = \frac{1}{N}\sum_{t=1}^{N}[a1(t)a2(t) + b1(t)b2(t) + a1(t)b2(t) + b1(t)a2(t) + \\ a3(t)a4(t) + b3(t)b4(t) + a3(t)b4(t) + b3(t)a4(t)]$$

where $ai(t)$ and $bi(t)$, $i \in \{1,2,3,4\}$ are signals generated by filtering the two mutually independent Gaussian white noise signals $e(t)$ and $f(t)$ through the stable scalar filters $A1, B1, A3, B3$ and the vectors of stable filters $A2, B2, A4, B4$.

$$\begin{array}{llll} a1(t) = A1e(t), & a2(t) = A2e(t), & a3(t) = A3e(t), & a4(t) = A4e(t) \\ b1(t) = B1f(t), & b2(t) = B2f(t), & b3(t) = B3f(t), & b4(t) = B4f(t) \end{array}$$

Using (A.2) *and evaluating the terms yields the following sum of cross correlation functions:*

$$Cov[Q_N] = \frac{1}{N^2}\sum_{t,s=1}^{N} R_{a1a1}(t-s)R_{a2a2^T}(t-s) + R_{a2a1}(t-s)R_{a1a2^T}(t-s) + \\ R_{a1a3}(t-s)R_{a2a4^T}(t-s) + R_{a2a3}(t-s)R_{a1a4^T}(t-s) + R_{b1b1}(t-s)R_{b2b2^T}(t-s) + \\ R_{b2b1}(t-s)R_{b1b2^T}(t-s) + R_{b1b3}(t-s)R_{b2b4^T}(t-s) + R_{b2b3}(t-s)R_{b1b4^T}(t-s) + \\ R_{a3a1}(t-s)R_{a4a2^T}(t-s) + R_{a4a1}(t-s)R_{a3a2^T}(t-s) + R_{a3a3}(t-s)R_{a4a4^T}(t-s) + \\ R_{a4a3}(t-s)R_{a3a4^T}(t-s) + R_{b3b1}(t-s)R_{b4b2^T}(t-s) + R_{b4b1}(t-s)R_{b3b2^T}(t-s) + \\ R_{b3b3}(t-s)R_{b4b4^T}(t-s) + R_{b4b3}(t-s)R_{b3b4^T}(t-s) + R_{a1a1}(t-s)R_{b2b2^T}(t-s) + \\ R_{a1a2}(t-s)R_{b2b1^T}(t-s) + R_{a1a3}(t-s)R_{b2b4^T}(t-s) + R_{a1a4}(t-s)R_{b2b3^T}(t-s) + \\ R_{a2a1}(t-s)R_{b1b2^T}(t-s) + R_{a2a2}(t-s)R_{b1b1^T}(t-s) + R_{a2a3}(t-s)R_{b1b4^T}(t-s) + \\ R_{a2a4}(t-s)R_{b1b3^T}(t-s) + R_{a3a1}(t-s)R_{b4b2^T}(t-s) + R_{a3a2}(t-s)R_{b4b1^T}(t-s) + \\ R_{a3a3}(t-s)R_{b4b4^T}(t-s) + R_{a3a4}(t-s)R_{b4b3^T}(t-s) + R_{a4a1}(t-s)R_{b3b2^T}(t-s) + \\ R_{a4a2}(t-s)R_{b3b1^T}(t-s) + R_{a4a3}(t-s)R_{b3b4^T}(t-s) + R_{a4a4}(t-s)R_{b3b3^T}(t-s)$$

Hence

$$
\begin{aligned}
\lim_{N\to\infty} N\,Cov[Q_N] =& \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_{a1a1}(\omega)\overline{\Phi}_{a2a2^T}(\omega) + \Phi_{a2a1}(\omega)\overline{\Phi}_{a1a2^T}(\omega) + \\
& \Phi_{a1a3}(\omega)\overline{\Phi}_{a2a4^T}(\omega) + \Phi_{a2a3}(\omega)\overline{\Phi}_{a1a4^T}(\omega) + \Phi_{b1b1}(\omega)\overline{\Phi}_{b2b2^T}(\omega) + \\
& \Phi_{b2b1}(\omega)\overline{\Phi}_{b1b2^T}(\omega) + \Phi_{b1b3}(\omega)\overline{\Phi}_{b2b4^T}(\omega) + \Phi_{b2b3}(\omega)\overline{\Phi}_{b1b4^T}(\omega) + \\
& \Phi_{a3a1}(\omega)\overline{\Phi}_{a4a2^T}(\omega) + \Phi_{a4a1}(\omega)\overline{\Phi}_{a3a2^T}(\omega) + \Phi_{a3a3}(\omega)\overline{\Phi}_{a4a4^T}(\omega) + \\
& \Phi_{a4a3}(\omega)\overline{\Phi}_{a3a4^T}(\omega) + \Phi_{b3b1}(\omega)\overline{\Phi}_{b4b2^T}(\omega) + \Phi_{b4b1}(\omega)\overline{\Phi}_{b3b2^T}(\omega) + \\
& \Phi_{b3b3}(\omega)\overline{\Phi}_{b4b4^T}(\omega) + \Phi_{b4b3}(\omega)\overline{\Phi}_{b3b4^T}(\omega) + \Phi_{a1a1}(\omega)\overline{\Phi}_{b2b2^T}(\omega) + \\
& \Phi_{a1a2}(\omega)\overline{\Phi}_{b2b1^T}(\omega) + \Phi_{a1a3}(\omega)\overline{\Phi}_{b2b4^T}(\omega) + \Phi_{a1a4}(\omega)\overline{\Phi}_{b2b3^T}(\omega) + \\
& \Phi_{a2a1}(\omega)\overline{\Phi}_{b1b2^T}(\omega) + \Phi_{a2a2}(\omega)\overline{\Phi}_{b1b1^T}(\omega) + \Phi_{a2a3}(\omega)\overline{\Phi}_{b1b4^T}(\omega) + \\
& \Phi_{a2a4}(\omega)\overline{\Phi}_{b1b3^T}(\omega) + \Phi_{a3a1}(\omega)\overline{\Phi}_{b4b2^T}(\omega) + \Phi_{a3a2}(\omega)\overline{\Phi}_{b4b1^T}(\omega) + \\
& \Phi_{a3a3}(\omega)\overline{\Phi}_{b4b4^T}(\omega) + \Phi_{a3a4}(\omega)\overline{\Phi}_{b4b3^T}(\omega) + \Phi_{a4a1}(\omega)\overline{\Phi}_{b3b2^T}(\omega) + \\
& \Phi_{a4a2}(\omega)\overline{\Phi}_{b3b1^T}(\omega) + \Phi_{a4a3}(\omega)\overline{\Phi}_{b3b4^T}(\omega) + \Phi_{a4a4}(\omega)\overline{\Phi}_{b3b3^T}(\omega)d\omega \qquad \text{(B.1)}
\end{aligned}
$$

When $Ai, Bi,\ i \in \{1,2,3,4\}$ refers to transfer functions in Equation (3.1) the cross spectra are

$$
\begin{aligned}
\Phi_{a1a1} =& |S(e^{j\omega},\boldsymbol{\rho})|^2|G(e^{j\omega})|^2\Phi_{r^p} \\
\Phi_{a2a2} =& |S(e^{j\omega},\boldsymbol{\rho})|^4|G(e^{j\omega})|^4\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\frac{\partial C^*(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\Phi_{r^p} \\
\Phi_{a3a3} =& \lambda|S(e^{j\omega},\boldsymbol{\rho})|^2\Phi_{r^p} \\
\Phi_{a4a4} =& \lambda|S(e^{j\omega},\boldsymbol{\rho})|^4|G(e^{j\omega})|^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\frac{\partial C^*(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\Phi_{r^p} \\
\Phi_{a1a2} =& -|S(e^{j\omega},\boldsymbol{\rho})|^2|G(e^{j\omega})|^2\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}S(e^{j\omega},\boldsymbol{\rho})G(e^{j\omega})\right)^*\Phi_{r^p} \\
\Phi_{a2a1} =& -|S(e^{j\omega},\boldsymbol{\rho})|^2|G(e^{j\omega})|^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}S(e^{j\omega},\boldsymbol{\rho})G(e^{j\omega})\Phi_{r^p} \\
\Phi_{a1a3} =& \sqrt{\lambda}|S(e^{j\omega},\boldsymbol{\rho})|^2G(e^{j\omega})\Phi_{r^p} \\
\Phi_{a3a1} =& \sqrt{\lambda}|S(e^{j\omega},\boldsymbol{\rho})|^2\overline{G}(e^{j\omega})\Phi_{r^p} \\
\Phi_{a1a4} =& -\sqrt{\lambda}|S(e^{j\omega},\boldsymbol{\rho})|^2|G(e^{j\omega})|^2\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}S(e^{j\omega},\boldsymbol{\rho})\right)^*\Phi_{r^p} \\
\Phi_{a4a1} =& -\sqrt{\lambda}|S(e^{j\omega},\boldsymbol{\rho})|^2|G(e^{j\omega})|^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}S(e^{j\omega},\boldsymbol{\rho})\Phi_{r^p} \\
\Phi_{a2a3} =& -\sqrt{\lambda}|S(e^{j\omega},\boldsymbol{\rho})|^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}G(e^{j\omega})^2S(e^{j\omega},\boldsymbol{\rho})\Phi_{r^p} \\
\Phi_{a3a2} =& -\sqrt{\lambda}|S(e^{j\omega},\boldsymbol{\rho})|^2\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}G(e^{j\omega})^2S(e^{j\omega},\boldsymbol{\rho})\right)^*\Phi_{r^p} \\
\Phi_{a2a4} =& \sqrt{\lambda}|S(e^{j\omega},\boldsymbol{\rho})|^4|G(e^{j\omega})|^2G(e^{j\omega})\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\frac{\partial C^*(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\Phi_{r^p} \\
\Phi_{a4a2} =& \sqrt{\lambda}|S(e^{j\omega},\boldsymbol{\rho})|^4|G(e^{j\omega})|^2\overline{G}(e^{j\omega})\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\frac{\partial C^*(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\Phi_{r^p} \\
\Phi_{a3a4} =& -\lambda|S(e^{j\omega},\boldsymbol{\rho})|^2\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}S(e^{j\omega},\boldsymbol{\rho})G(e^{j\omega})\right)^*\Phi_{r^p} \\
\Phi_{a4a3} =& -\lambda|S(e^{j\omega},\boldsymbol{\rho})|^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}S(e^{j\omega},\boldsymbol{\rho})G(e^{j\omega})\Phi_{r^p}
\end{aligned}
$$

$$\begin{aligned}
\Phi_{b1b1} =& |S(e^{j\omega},\boldsymbol{\rho})|^2\Phi_v \\
\Phi_{b2b2} =& |S(e^{j\omega},\boldsymbol{\rho})|^4|G(e^{j\omega})|^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\frac{\partial C^*(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\Phi_v \\
\Phi_{b3b3} =& \lambda|S(e^{j\omega},\boldsymbol{\rho})|^2|C(e^{j\omega},\boldsymbol{\rho})|^2\Phi_v \\
\Phi_{b4b4} =& \lambda|S(e^{j\omega},\boldsymbol{\rho})|^4\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\frac{\partial C^*(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\Phi_v \\
\Phi_{b1b2} =& -|S(e^{j\omega},\boldsymbol{\rho})|^2\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}S(e^{j\omega},\boldsymbol{\rho})G(e^{j\omega})\right)^*\Phi_v \\
\Phi_{b2b1} =& -|S(e^{j\omega},\boldsymbol{\rho})|^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}S(e^{j\omega},\boldsymbol{\rho})G(e^{j\omega})\Phi_v \\
\Phi_{b1b3} =& -\sqrt{\lambda}|S(e^{j\omega},\boldsymbol{\rho})|^2\overline{C}(e^{j\omega},\boldsymbol{\rho})\Phi_v \\
\Phi_{b3b1} =& -\sqrt{\lambda}|S(e^{j\omega},\boldsymbol{\rho})|^2C(e^{j\omega},\boldsymbol{\rho})\Phi_v \\
\Phi_{b1b4} =& -\sqrt{\lambda}|S(e^{j\omega},\boldsymbol{\rho})|^2\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}S(e^{j\omega},\boldsymbol{\rho})\right)^*\Phi_v \\
\Phi_{b4b1} =& -\sqrt{\lambda}|S(e^{j\omega},\boldsymbol{\rho})|^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}S(e^{j\omega},\boldsymbol{\rho})\Phi_v \\
\Phi_{b2b3} =& \sqrt{\lambda}|S(e^{j\omega},\boldsymbol{\rho})|^2\overline{C}(e^{j\omega},\boldsymbol{\rho})\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}G(e^{j\omega})^2S(e^{j\omega},\boldsymbol{\rho})\Phi_v \\
\Phi_{b3b2} =& \sqrt{\lambda}|S(e^{j\omega},\boldsymbol{\rho})|^2C(e^{j\omega},\boldsymbol{\rho})\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}G(e^{j\omega})^2S(e^{j\omega},\boldsymbol{\rho})\right)^*\Phi_v \\
\Phi_{b2b4} =& \sqrt{\lambda}|S(e^{j\omega},\boldsymbol{\rho})|^4G(e^{j\omega})\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\frac{\partial C^*(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\Phi_v \\
\Phi_{b4b2} =& \sqrt{\lambda}|S(e^{j\omega},\boldsymbol{\rho})|^4\overline{G}(e^{j\omega})\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\frac{\partial C^*(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\Phi_v \\
\Phi_{b3b4} =& \lambda|S(e^{j\omega},\boldsymbol{\rho})|^2C(e^{j\omega},\boldsymbol{\rho})\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}S(e^{j\omega},\boldsymbol{\rho})\right)^*\Phi_v \\
\Phi_{b4b3} =& \lambda|S(e^{j\omega},\boldsymbol{\rho})|^2\overline{C}(e^{j\omega},\boldsymbol{\rho})\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}S(e^{j\omega},\boldsymbol{\rho})\Phi_v
\end{aligned}$$

Evaluating the multiplication of the cross spectra in (B.1), *it is evident that the terms can be divided into four groups. The terms in these four sub-groups are evaluated separately and summed.*

In the following two complex functions are utilized

$$\Psi(e^{j\omega},\boldsymbol{\rho}) =[G(e^{j\omega},\boldsymbol{\rho})-\lambda\overline{C}(e^{j\omega},\boldsymbol{\rho})]S(e^{j\omega},\boldsymbol{\rho})\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}} \tag{B.2}$$

$$\Upsilon(e^{j\omega},\boldsymbol{\rho}) =[|G(e^{j\omega},\boldsymbol{\rho})|^2+\lambda]S(e^{j\omega},\boldsymbol{\rho})\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}} \tag{B.3}$$

$$
\begin{aligned}
\Phi_{a1a1}\overline{\Phi}_{a2a2^T} =& |S(e^{j\omega},\boldsymbol{\rho})|^6|G(e^{j\omega})|^6\overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\right)^T\Phi_{r^p}^2\\
\Phi_{a2a1}\overline{\Phi}_{a1a2^T} =& |S(e^{j\omega},\boldsymbol{\rho})|^4|G(e^{j\omega})|^4S(e^{j\omega},\boldsymbol{\rho})^2G(e^{j\omega})^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\right)^T\Phi_{r^p}^2\\
\Phi_{a1a3}\overline{\Phi}_{a2a4^T} =& \lambda|S(e^{j\omega},\boldsymbol{\rho})|^6|G(e^{j\omega})|^4\overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\right)^T\Phi_{r^p}^2\\
\Phi_{a2a3}\overline{\Phi}_{a1a4^T} =& \lambda|S(e^{j\omega},\boldsymbol{\rho})|^4|G(e^{j\omega})|^2S(e^{j\omega},\boldsymbol{\rho})^2G(e^{j\omega})^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\right)^T\Phi_{r^p}^2\\
\Phi_{a3a1}\overline{\Phi}_{a4a2^T} =& \lambda|S(e^{j\omega},\boldsymbol{\rho})|^6|G(e^{j\omega})|^4\overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\right)^T\Phi_{r^p}^2\\
\Phi_{a4a1}\overline{\Phi}_{a3a2^T} =& \lambda|S(e^{j\omega},\boldsymbol{\rho})|^4|G(e^{j\omega})|^2S(e^{j\omega},\boldsymbol{\rho})^2G(e^{j\omega})^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\right)^T\Phi_{r^p}^2\\
\Phi_{a3a3}\overline{\Phi}_{a4a4^T} =& \lambda^2|S(e^{j\omega},\boldsymbol{\rho})|^6|G(e^{j\omega})|^2\overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\right)^T\Phi_{r^p}^2\\
\Phi_{a4a3}\overline{\Phi}_{a3a4^T} =& \lambda^2|S(e^{j\omega},\boldsymbol{\rho})|^4S(e^{j\omega},\boldsymbol{\rho})^2G(e^{j\omega})^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\right)^T\Phi_{r^p}^2
\end{aligned}
$$

the sum of which yields

$$
\begin{aligned}
\lim_{N\to\infty} N\,Cov[Q_N^1] =& \frac{1}{2\pi}\int_{-\pi}^{\pi}|S(e^{j\omega},\boldsymbol{\rho})|^4\Phi_{r^p}^2\times\\
&\left[G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho})(G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho}))^T+\overline{(G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho}))}(G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho}))^T\right]d\omega\\
=& \frac{2}{2\pi}\int_{-\pi}^{\pi}|S(e^{j\omega},\boldsymbol{\rho})|^4\left[\mathcal{R}e\{G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho})\}\mathcal{R}e\{G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho})\}^T\right]\Phi_{r^p}^2d\omega
\end{aligned}
$$

$$
\begin{aligned}
\Phi_{b1b1}\overline{\Phi}_{b2b2^T} =& |S(e^{j\omega},\boldsymbol{\rho})|^6|G(e^{j\omega})|^2\overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\right)^T\Phi_v^2\\
\Phi_{b2b1}\overline{\Phi}_{b1b2^T} =& |S(e^{j\omega},\boldsymbol{\rho})|^4S(e^{j\omega},\boldsymbol{\rho})^2G(e^{j\omega})^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\right)^T\Phi_v^2\\
\Phi_{b1b3}\overline{\Phi}_{b2b4^T} =& -\lambda|S(e^{j\omega},\boldsymbol{\rho})|^6\overline{G}(e^{j\omega})\overline{C}(e^{j\omega},\boldsymbol{\rho})\overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\right)^T\Phi_v^2\\
\Phi_{b2b3}\overline{\Phi}_{b1b4^T} =& -\lambda|S(e^{j\omega},\boldsymbol{\rho})|^4G(e^{j\omega})S(e^{j\omega},\boldsymbol{\rho})^2\overline{C}(e^{j\omega},\boldsymbol{\rho})\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\right)^T\Phi_v^2\\
\Phi_{b3b1}\overline{\Phi}_{b4b2^T} =& -\lambda|S(e^{j\omega},\boldsymbol{\rho})|^6G(e^{j\omega})C(e^{j\omega},\boldsymbol{\rho})\overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\right)^T\Phi_v^2\\
\Phi_{b4b1}\overline{\Phi}_{b3b2^T} =& -\lambda|S(e^{j\omega},\boldsymbol{\rho})|^4G(e^{j\omega})S(e^{j\omega},\boldsymbol{\rho})^2\overline{C}(e^{j\omega},\boldsymbol{\rho})\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\right)^T\Phi_v^2\\
\Phi_{b3b3}\overline{\Phi}_{b4b4^T} =& \lambda^2|S(e^{j\omega},\boldsymbol{\rho})|^6|C(e^{j\omega},\boldsymbol{\rho})|^2\overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\right)^T\Phi_v^2\\
\Phi_{b4b3}\overline{\Phi}_{b3b4^T} =& \lambda^2|S(e^{j\omega},\boldsymbol{\rho})|^4S(e^{j\omega},\boldsymbol{\rho})^2\overline{C}(e^{j\omega},\boldsymbol{\rho})^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\right)^T\Phi_v^2
\end{aligned}
$$

the sum of which yields

$$
\begin{aligned}
\lim_{N\to\infty} N\,Cov[Q_N^2] =& \frac{1}{2\pi}\int_{-\pi}^{\pi} |S(e^{j\omega},\boldsymbol{\rho})|^4 \Phi_v^2 \times \\
& \left[\Psi(e^{j\omega},\boldsymbol{\rho})\Psi(e^{j\omega},\boldsymbol{\rho})^T + \overline{\Psi}(e^{j\omega},\boldsymbol{\rho})\Psi(e^{j\omega},\boldsymbol{\rho})^T\right] d\omega \\
=& \frac{2}{2\pi}\int_{-\pi}^{\pi} |S(e^{j\omega},\boldsymbol{\rho})|^4 \left[\mathcal{R}e\{\Psi(e^{j\omega},\boldsymbol{\rho})\}\mathcal{R}e\{\Psi(e^{j\omega},\boldsymbol{\rho})\}^T\right] \Phi_v^2 d\omega
\end{aligned}
$$

$$
\begin{aligned}
\Phi_{a1a1}\overline{\Phi}_{b2b2^T} =& |S(e^{j\omega},\boldsymbol{\rho})|^6 |G(e^{j\omega})|^4 \overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}} \left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\right)^T \Phi_{r^p}\Phi_v \\
\Phi_{a1a2}\overline{\Phi}_{b2b1^T} =& |S(e^{j\omega},\boldsymbol{\rho})|^4 |G(e^{j\omega})|^2 (\overline{S}(e^{j\omega},\boldsymbol{\rho})\overline{G}(e^{j\omega}))^2 \frac{\partial C^*(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}} \overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}} \Phi_{r^p}\Phi_v \\
\Phi_{a1a3}\overline{\Phi}_{b2b4^T} =& \lambda |S(e^{j\omega},\boldsymbol{\rho})|^6 |G(e^{j\omega})|^2 \overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}} \left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\right)^T \Phi_{r^p}\Phi_v \\
\Phi_{a1a4}\overline{\Phi}_{b2b3^T} =& -\lambda |S(e^{j\omega},\boldsymbol{\rho})|^4 |G(e^{j\omega})|^2 \overline{S}(e^{j\omega},\boldsymbol{\rho})^2 \overline{G}(e^{j\omega}) C(e^{j\omega},\boldsymbol{\rho}) \frac{\partial C^*(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}} \overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}} \Phi_{r^p}\Phi_v \\
\Phi_{a3a1}\overline{\Phi}_{b4b2^T} =& \lambda |S(e^{j\omega},\boldsymbol{\rho})|^6 |G(e^{j\omega})|^2 \overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}} \left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\right)^T \Phi_{r^p}\Phi_v \\
\Phi_{a3a2}\overline{\Phi}_{b4b1^T} =& \lambda |S(e^{j\omega},\boldsymbol{\rho})|^4 (\overline{S}(e^{j\omega},\boldsymbol{\rho})\overline{G}(e^{j\omega}))^2 \frac{\partial C^*(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}} \overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}} \Phi_{r^p}\Phi_v \\
\Phi_{a3a3}\overline{\Phi}_{b4b4^T} =& \lambda^2 |S(e^{j\omega},\boldsymbol{\rho})|^6 \overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}} \left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}\right)^T \Phi_{r^p}\Phi_v \\
\Phi_{a3a4}\overline{\Phi}_{b4b3^T} =& -\lambda^2 |S(e^{j\omega},\boldsymbol{\rho})|^4 \overline{S}(e^{j\omega},\boldsymbol{\rho})^2 \overline{G}(e^{j\omega}) C(e^{j\omega},\boldsymbol{\rho}) \frac{\partial C^*(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}} \overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial \boldsymbol{\rho}}} \Phi_{r^p}\Phi_v
\end{aligned}
$$

the sum of which yields

$$
\begin{aligned}
\lim_{N\to\infty} N\,Cov[Q_N^3] =& \frac{1}{2\pi}\int_{-\pi}^{\pi} |S(e^{j\omega},\boldsymbol{\rho})|^4 \Phi_{r^p}\Phi_v \times \\
& \left[G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho})(\Psi(e^{j\omega},\boldsymbol{\rho}))^T + \overline{\Upsilon}(e^{j\omega},\boldsymbol{\rho})\Upsilon(e^{j\omega},\boldsymbol{\rho})^T\right] d\omega \\
=& \frac{1}{2\pi}\int_{-\pi}^{\pi} |S(e^{j\omega},\boldsymbol{\rho})|^4 \times \\
& \Big[\mathcal{R}e\{G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho})\}\mathcal{R}e\{\Psi(e^{j\omega},\boldsymbol{\rho})\}^T + \\
& \mathcal{I}m\{G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho})\}\mathcal{I}m\{\Psi(e^{j\omega},\boldsymbol{\rho})\}^T + \\
& \mathcal{R}e\{\Upsilon(e^{j\omega},\boldsymbol{\rho})\}\mathcal{R}e\{\Upsilon(e^{j\omega},\boldsymbol{\rho})\}^T + \\
& \mathcal{I}m\{\Upsilon(e^{j\omega},\boldsymbol{\rho})\}\mathcal{I}m\{\Upsilon(e^{j\omega},\boldsymbol{\rho})\}^T\Big] \Phi_{r^p}\Phi_v d\omega
\end{aligned}
$$

$$
\begin{aligned}
\Phi_{a2a1}\overline{\Phi}_{b1b2^T} =& |S(e^{j\omega},\boldsymbol{\rho})|^4|G(e^{j\omega})|^2(S(e^{j\omega},\boldsymbol{\rho})G(e^{j\omega}))^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T\Phi_{r^p}\Phi_v\\
\Phi_{a2a2}\overline{\Phi}_{b1b1^T} =& |S(e^{j\omega},\boldsymbol{\rho})|^6|G(e^{j\omega})|^2(\overline{S}(e^{j\omega},\boldsymbol{\rho})\overline{G}(e^{j\omega}))^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\frac{\partial C^*(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\Phi_{r^p}\Phi_v\\
\Phi_{a2a3}\overline{\Phi}_{b1b4^T} =& \lambda|S(e^{j\omega},\boldsymbol{\rho})|^4|G(e^{j\omega})|^2(S(e^{j\omega},\boldsymbol{\rho})G(e^{j\omega}))^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T\Phi_{r^p}\Phi_v\\
\Phi_{a2a4}\overline{\Phi}_{b1b3^T} =& -\lambda|S(e^{j\omega},\boldsymbol{\rho})|^6|G(e^{j\omega})|^2S(e^{j\omega},\boldsymbol{\rho})C(e^{j\omega,\boldsymbol{\rho}})\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\frac{\partial C^*(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\Phi_{r^p}\Phi_v\\
\Phi_{a4a1}\overline{\Phi}_{b3b2^T} =& -\lambda|S(e^{j\omega},\boldsymbol{\rho})|^4|G(e^{j\omega})|^2S(e^{j\omega},\boldsymbol{\rho})^2G(e^{j\omega})\overline{C}(e^{j\omega},\boldsymbol{\rho})\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T\Phi_{r^p}\Phi_v\\
\Phi_{a4a2}\overline{\Phi}_{b3b1^T} =& -\lambda|S(e^{j\omega},\boldsymbol{\rho})|^6|G(e^{j\omega})|^2\overline{G}(e^{j\omega}))^2\overline{C}(e^{j\omega},\boldsymbol{\rho})\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\frac{\partial C^*(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\Phi_{r^p}\Phi_v\\
\Phi_{a4a3}\overline{\Phi}_{b3b4^T} =& -\lambda^2|S(e^{j\omega},\boldsymbol{\rho})|^4S(e^{j\omega},\boldsymbol{\rho})^2G(e^{j\omega})\overline{C}(e^{j\omega},\boldsymbol{\rho})\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T\Phi_{r^p}\Phi_v\\
\Phi_{a4a4}\overline{\Phi}_{b3b3^T} =& \lambda^2|S(e^{j\omega},\boldsymbol{\rho})|^6|G(e^{j\omega})|^2|C(e^{j\omega},\boldsymbol{\rho})|^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\frac{\partial C^*(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\Phi_{r^p}\Phi_v
\end{aligned}
$$

the sum of which yields

$$
\begin{aligned}
\lim_{N\to\infty} N\,Cov[Q_N^4] =& \frac{1}{2\pi}\int_{-\pi}^{\pi}|S(e^{j\omega},\boldsymbol{\rho})|^4\Phi_{r^p}\Phi_v\times\\
&\left[G(e^{j\omega})\Psi(e^{j\omega},\boldsymbol{\rho})(G(e^{j\omega})\Psi(e^{j\omega},\boldsymbol{\rho}))^* + (G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho}))^*\overline{\Psi}(e^{j\omega},\boldsymbol{\rho})\right]d\omega\\
=& \frac{1}{2\pi}\int_{-\pi}^{\pi}|S(e^{j\omega},\boldsymbol{\rho})|^4\times\\
&\Big[\mathcal{R}e\{G(e^{j\omega})\Psi(e^{j\omega},\boldsymbol{\rho})\}\mathcal{R}e\{G(e^{j\omega})\Psi(e^{j\omega},\boldsymbol{\rho})\}^T+\\
&\mathcal{I}m\{G(e^{j\omega})\Psi(e^{j\omega},\boldsymbol{\rho})\}\mathcal{I}m\{G(e^{j\omega})\Psi(e^{j\omega},\boldsymbol{\rho})\}^T+\\
&\mathcal{R}e\{G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho})\}^T\mathcal{R}e\{\Psi(e^{j\omega},\boldsymbol{\rho})\}-\\
&\mathcal{I}m\{G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho})\}^T\mathcal{I}m\{\Psi(e^{j\omega},\boldsymbol{\rho})\}\Big]\Phi_{r^p}\Phi_v d\omega
\end{aligned}
$$

Combining these four terms gives the covariance expression for S_N.

$$
\begin{aligned}
\lim_{N\to\infty} N\,Cov[S_N(\boldsymbol{\rho})] = & \frac{1}{2\pi}\int_{-\pi}^{\pi}|S(e^{j\omega},\boldsymbol{\rho})|^4\times[\\
&\left[G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho})(G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho}))^T + \overline{(G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho}))}(G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho}))^T\right]\Phi_{r^p}^2+\\
&\left[\Psi(e^{j\omega},\boldsymbol{\rho})\Psi(e^{j\omega},\boldsymbol{\rho})^T + \overline{\Psi}(e^{j\omega},\boldsymbol{\rho})\Psi(e^{j\omega},\boldsymbol{\rho})^T\right]\Phi_v^2+\\
&\left[G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho})(\Psi(e^{j\omega},\boldsymbol{\rho}))^T + \overline{\Upsilon}(e^{j\omega},\boldsymbol{\rho})\Upsilon(e^{j\omega},\boldsymbol{\rho})^T\right]\Phi_{r^p}\Phi_v+\\
&\left[G(e^{j\omega})\Psi(e^{j\omega},\boldsymbol{\rho})(G(e^{j\omega})\Psi(e^{j\omega},\boldsymbol{\rho}))^* + (G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho}))^*\overline{\Psi}(e^{j\omega},\boldsymbol{\rho})\right]\Phi_{r^p}\Phi_v\ ]\,d\omega
\end{aligned}
$$

or

$$\lim_{N\to\infty} N\,Cov[S_N(\boldsymbol{\rho})] = \frac{1}{2\pi}\int_{-\pi}^{\pi} |S(e^{j\omega},\boldsymbol{\rho})|^4 \times \Big[$$

$$\left[\mathcal{R}e\{G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho})\}\mathcal{R}e\{G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho})\}^T\right]\Phi_{r^p}^2 +$$
$$\left[\mathcal{R}e\{\Psi(e^{j\omega},\boldsymbol{\rho})\}\mathcal{R}e\{\Psi(e^{j\omega},\boldsymbol{\rho})\}^T\right]\Phi_v^2 +$$
$$\Big[2\mathcal{R}e\{G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho})\}\mathcal{R}e\{\Psi(e^{j\omega},\boldsymbol{\rho})\}^T + \mathcal{I}m\{G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho})\}\mathcal{I}m\{\Psi(e^{j\omega},\boldsymbol{\rho})\}^T -$$
$$\mathcal{I}m\{G(e^{j\omega})\Upsilon(e^{j\omega},\boldsymbol{\rho})\}^T\mathcal{I}m\{\Psi(e^{j\omega},\boldsymbol{\rho})\} + \mathcal{R}e\{G(e^{j\omega})\Psi(e^{j\omega},\boldsymbol{\rho})\}\mathcal{R}e\{G(e^{j\omega})\Psi(e^{j\omega},\boldsymbol{\rho})\}^T +$$
$$\mathcal{I}m\{G(e^{j\omega})\Psi(e^{j\omega},\boldsymbol{\rho})\}\mathcal{I}m\{G(e^{j\omega})\Psi(e^{j\omega},\boldsymbol{\rho})\}^T + \mathcal{R}e\{\Upsilon(e^{j\omega},\boldsymbol{\rho})\}\mathcal{R}e\{\Upsilon(e^{j\omega},\boldsymbol{\rho})\}^T +$$
$$\mathcal{I}m\{\Upsilon(e^{j\omega},\boldsymbol{\rho})\}\mathcal{I}m\{\Upsilon(e^{j\omega},\boldsymbol{\rho})\}^T\Big]\Phi_{r^p}\Phi_v\Big]d\omega$$

Which is the shortest possible representation of the asymptotic covariance of S_N from (3.1). **q.e.d.**

Derivation B.2 (Covariance expressions for E_N) *For the derivation of the covariance of E_N, let the sum Q_N be a generalization of the structure of E_N according to*

$$\begin{aligned} Q_N =& \frac{1}{N}\sum_{t=1}^{N}[(a1(t)+b1(t))c1(t)+(a2(t)+b2(t))c2(t)] \\ =& \frac{1}{N}\sum_{t=1}^{N}[a1(t)c1(t)+b1(t)c1(t)+a2(t)c2(t)+b2(t)c2(t)] \end{aligned}$$

where $ai(t)$, $bi(t)$ and $ci(t)$, $i \in \{1,2\}$ are signals generated by filtering the three mutually independent white noise signals $e(t)$, $f(t)$ and $g(t)$ through the stable scalar filters $A1, A2, B1, B2$ and the vectors of stable filters $C1, C2$.

$$\begin{aligned} a1(t) &= A1e(t), & a2(t) &= A2e(t) \\ b1(t) &= B1f(t), & b2(t) &= B2f(t) \\ c1(t) &= C1g(t), & c2(t) &= C2g(t) \end{aligned}$$

Using (A.2)*, evaluating the terms and realizing that the cross correlation function between independent signals is zero yields:*

$$\begin{aligned} Cov[Q_N] =& \frac{1}{N^2}\sum_{t,s=1}^{N} R_{a1a1}(t-s)R_{c1c1^T}(t-s) + R_{a1a2}(t-s)R_{c1c2^T}(t-s) + \\ & R_{b1b1}(t-s)R_{c1c1^T}(t-s) + R_{b1b2}(t-s)R_{c1c2^T}(t-s) + R_{a2a1}(t-s)R_{c2c1^T}(t-s) + \\ & R_{a2a2}(t-s)R_{c2c2^T}(t-s) + R_{b2b1}(t-s)R_{c2c1^T}(t-s) + R_{b2b2}(t-s)R_{c2c2^T}(t-s) \end{aligned}$$

Hence

$$\begin{aligned} \lim_{N\to\infty} N\,Cov[Q_N] =& \frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_{a1a1}(\omega)\overline{\Phi}_{c1c1^T}(\omega) + \Phi_{a1a2}(\omega)\overline{\Phi}_{c1c2^T}(\omega) + \\ & \Phi_{b1b1}(\omega)\overline{\Phi}_{c1c1^T}(\omega) + \Phi_{b1b2}(\omega)\overline{\Phi}_{c1c2^T}(\omega) + \Phi_{a2a1}(\omega)\overline{\Phi}_{c2c1^T}(\omega) + \\ & \Phi_{a2a2}(\omega)\overline{\Phi}_{c2c2^T}(\omega) + \Phi_{b2b1}(\omega)\overline{\Phi}_{c2c1^T}(\omega) + \Phi_{b2b2}(\omega)\overline{\Phi}_{c2c2^T}(\omega)d\omega \end{aligned} \tag{B.4}$$

When $A1, A2, B1, B2, C1$ and $C2$ refer to transfer functions in Equation (3.2) *the cross spectra* *are*

$$\begin{aligned}
\Phi_{a1a1} &= |S(e^{j\omega},\boldsymbol{\rho})|^2|G(e^{j\omega},\boldsymbol{\rho})|^2\Phi_{r^p}\\
\Phi_{a2a2} &= \lambda|S(e^{j\omega},\boldsymbol{\rho})|^2\Phi_{r^p}\\
\Phi_{a1a2} &= \sqrt{\lambda}G(e^{j\omega})|S(e^{j\omega},\boldsymbol{\rho})|^2\Phi_{r^p}\\
\Phi_{a2a1} &= \sqrt{\lambda}G^*(e^{j\omega})|S(e^{j\omega},\boldsymbol{\rho})|^2\Phi_{r^p}\\
\Phi_{b1b1} &= |S(e^{j\omega},\boldsymbol{\rho})|^2\Phi_v\\
\Phi_{b1b2} &= -\sqrt{\lambda}C^*(e^{j\omega},\boldsymbol{\rho})|S(e^{j\omega},\boldsymbol{\rho})|^2\Phi_v\\
\Phi_{b2b2} &= \lambda|S(e^{j\omega},\boldsymbol{\rho})|^2|C(e^{j\omega},\boldsymbol{\rho})|^2\Phi_v\\
\Phi_{b2b1} &= -\sqrt{\lambda}C(e^{j\omega},\boldsymbol{\rho})|S(e^{j\omega},\boldsymbol{\rho})|^2\Phi_v\\
\Phi_{c1c1^T} &= |S(e^{j\omega},\boldsymbol{\rho})|^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\frac{\partial C^*(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\Phi_v\\
\Phi_{c2c2^T} &= \lambda|C(e^{j\omega},\boldsymbol{\rho})|^2|S(e^{j\omega},\boldsymbol{\rho})|^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\frac{\partial C^*(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\Phi_v\\
\Phi_{c1c2^T} &= -\sqrt{\lambda}C^*(e^{j\omega},\boldsymbol{\rho})|S(e^{j\omega},\boldsymbol{\rho})|^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\frac{\partial C^*(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\Phi_v\\
\Phi_{c2c1^T} &= -\sqrt{\lambda}C(e^{j\omega},\boldsymbol{\rho})|S(e^{j\omega},\boldsymbol{\rho})|^2\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\frac{\partial C^*(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\Phi_v
\end{aligned}$$

Hence

$$\begin{aligned}
\Phi_{a1a1}\overline{\Phi}_{c1c1^T} =&|G(e^{j\omega})|^2|S(e^{j\omega},\boldsymbol{\rho})|^4\overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T\Phi_{r^p}\Phi_v\\
\Phi_{a2a2}\overline{\Phi}_{c2c2^T} =&\lambda^2|C(e^{j\omega},\boldsymbol{\rho})|^2|S(e^{j\omega},\boldsymbol{\rho})|^2\overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T\Phi_{r^p}\Phi_v\\
\Phi_{a1a2}\overline{\Phi}_{c2c1^T} =&-\lambda G(e^{j\omega})C(e^{j\omega},\boldsymbol{\rho})|S(e^{j\omega},\boldsymbol{\rho})|^4\overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T\Phi_{r^p}\Phi_v\\
\Phi_{a2a1}\overline{\Phi}_{c1c2^T} =&-\lambda G^*(e^{j\omega})\overline{C}(e^{j\omega},\boldsymbol{\rho})|S(e^{j\omega},\boldsymbol{\rho})|^4\overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T\Phi_{r^p}\Phi_v\\
\Phi_{b1b1}\overline{\Phi}_{c1c1^T} =&|S(e^{j\omega},\boldsymbol{\rho})|^4\overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T\Phi_v^2\\
\Phi_{b2b2}\overline{\Phi}_{c2c2^T} =&\lambda^2|C(e^{j\omega},\boldsymbol{\rho})|^4|S(e^{j\omega},\boldsymbol{\rho})|^4\overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T\Phi_v^2\\
\Phi_{b1b2}\overline{\Phi}_{c2c1^T} =&\lambda|C(e^{j\omega},\boldsymbol{\rho})|^2|S(e^{j\omega},\boldsymbol{\rho})|^4\overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T\Phi_v^2\\
\Phi_{b2b1}\overline{\Phi}_{c1c2^T} =&\Phi_{b1b2}\overline{\Phi}_{c2c1^T}
\end{aligned}$$

Inserting these expressions in (B.4) *gives:*

$$\begin{aligned}
\lim_{N\to\infty}N\,Cov[E_N(\boldsymbol{\rho})] =&\frac{1}{2\pi}\int_{-\pi}^{\pi}|S(e^{j\omega},\boldsymbol{\rho})|^4\left[1+\lambda|C(e^{j\omega},\boldsymbol{\rho})|^2\right]^2\overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T\Phi_v^2+\\
&|S(e^{j\omega},\boldsymbol{\rho})|^4\left[\overline{\Psi(e^{j\omega},\boldsymbol{\rho}))}\Psi(e^{j\omega},\boldsymbol{\rho}))^T\right]\Phi_{r^p}\Phi_v d\omega\\
\lim_{N\to\infty}N\,Cov[E_N(\boldsymbol{\rho})] =&\frac{1}{2\pi}\int_{-\pi}^{\pi}|S(e^{j\omega},\boldsymbol{\rho})|^4\left[1+\lambda|C(e^{j\omega},\boldsymbol{\rho})|^2\right]^2\overline{\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}}\left(\frac{\partial C(e^{j\omega},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}}\right)^T\Phi_v^2+\\
&\left[\mathcal{R}e\{\Psi(e^{j\omega},\boldsymbol{\rho}))\}\mathcal{R}e\{\Psi(e^{j\omega},\boldsymbol{\rho}))\}^T+\mathcal{I}m\{\Psi(e^{j\omega},\boldsymbol{\rho}))\}\mathcal{I}m\{\Psi(e^{j\omega},\boldsymbol{\rho}))\}^T\right]\times\\
&|S(e^{j\omega},\boldsymbol{\rho})|^4\Phi_{r^p}\Phi_v d\omega
\end{aligned}$$

This is the shortest possible representation of the asymptotic covariance of E_N from (3.2). **q.e.d.**

References

Åström, K. J. (1970). *Introduction to Stochastic Control Theory*. Academic Press.

Åström, K. J. and Hägglund, T. (1995). *PID Controllers: Theory, Design and Tuning*. Instrument Society of America, Research Triangle Park, North Carolina, second edition.

Bombois, X.; Scorletti, G.; Gevers, M.; Hildebrand, R. and van den Hof, P. (2004). Cheapest open-loop identification for control. In *Proceedings of the 34rd IEEE Conference on Decision and Control*, pages 382 – 387.

Box, G. E. P. and Jenkins, G. M. (1970). *Time Series Analysis forecasting and control.*. Holden-Day.

Gevers, M. (2002). A decade of progress in iterative process control design: from theory to practice. *Journal of process control*, **12**(4), 519–531.

Gevers, M. and Ljung, L. (1986). Optimal Experiment Design with Respect to the Intended Model Application. *Automatica*, **22**(5), 543–554.

Goodwin, G. C. and Payne, R. L. (1977). *Dynamic System Identification: Experiment Design and Data Analysis*, volume 136 of *Mathematics in science and engineering*. Academic Press.

Hildebrand, R.; Lecchini, A.; Solari, G. and Gevers, M. (2003). A Convergence Theorem for Iterative Feedback Tuning. Technical report, CESAME, Universit Catholique de Louvain, Louvain-la-Neuve, Belgium.

Hildebrand, R.; Lecchini, A.; Solari, G. and Gevers, M. (2005a). Asymptotic Accuracy of Iterative Feedback Tuning. *IEEE Transactions on Automatic Control*, **50**(8), 1182 – 1185.

Hildebrand, R.; Lecchini, A.; Solari, G. and Gevers, M. (2005b). Optimal Prefiltering in Iterative Feedback Tuning. *IEEE Transactions on Automatic Control*, **50**(8), 1196 – 1200.

Hjalmarsson, H. (2002). Iterative feedback tuning - an overview. *International journal of adaptive control and signal processing*, **16**, 373–395.

Hjalmarsson, H.; Gunnarsson, S. and Gevers, M. (1994). A Convergent Iterative Restricted Complexity Control Design Scheme. In *Proceedings of the 33rd IEEE Conference on Decision and Control*, volume 2, pages 1735 – 1740.

Hjalmarsson, H.; Gevers, M.; Gunnarsson, S. and Lequin, O. (1998). Iterative Feedback Tuning: Theory and Applications. *IEEE Control Systems Magazine*, **18**(4), 26–41.

Huusom, J. K.; Poulsen, N. K. and Jørgensen, S. B. (2008). Improving Convergence of Iterative Feedback Tuning. *In Press for Journal of Process Control.*. DOI information: 10.1016/j.jprocont.2008.09.004.

Jansson, H. and Hjalmarsson, H. (2005). Input Design via LMIs Admitting Frequency-wise Model Specifications in Confidence Regions. *IEEE Transactions on Automatic Control*, **50**(10), 1534–1549.

Robbins, H. and Monro, S. (1951). A stochastic approximation method. *Annals of mathematical statistics*, **22**(3), 400–407.

Solari, G. and Gevers, M. (2004). Unbiased Estimation of the Hessian for Iterative Feedback Tuning (IFT). In *Proceedings of 43rd IEEE Conference on Decision and Control*, volume 2, pages 1759–1760.